An NMR investigation of the structure-function characteristics of the *Pseudomonas aeruginosa* dimethylarginine dimethylaminohydrolase

Beatriz Simas Magalhães

University College London

A thesis submitted for the degree of Doctor of Philosophy

May 2006

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Thesis abstract

The enzyme dimethylarginine dimethylaminohydrolase (DDAH) is responsible for the metabolism of free asymmetric methylarginine residues (AMAs) to citrulline and mono- or di-methylamines. In mammals, the asymmetric methylarginines N^{η},N^{η}-dimethylarginine (ADMA) and N^{η}-methylarginine (L-NMMA) are endogenous inhibitors of all nitric oxide synthase (NOS) isoforms. By controlling local levels of AMAs, DDAH activity is implicated in the regulation of nitric oxide (NO) production, which gives rise to interest in the therapeutic potential of this enzyme.

This work describes the application of NMR methodology to study the backbone dynamics of *Pa*DDAH and interactions with ligands and inhibitors, as this protein provides an attractive template to further investigate the structure-function characteristics of DDAH isoforms and β/α-propeller enzymes in general.

To counter the relatively large size of the enzyme for NMR studies and the difficulties to obtain unambiguous resonance assignments for the amide backbone resonances of the 58 kD homodimer, a monomeric and fully active variant of *Pa*DDAH was engineered through rational design of site-directed mutation of two interfacial residues. The exclusively monomeric R40E-R98H double mutant protein proved more tractable by heteronuclear NMR than the wild-type homodimer. ^{15}N backbone relaxation studies of the mutant protein, analysed in the model-free formalism, revealed that the loop that closes down on the ligand in the active site displays low values of the generalised N-H bond order parameter (S^2) consistent with a high degree of mobility on the pico- to nanosecond timescale.

The interaction of *Pa*DDAH with a variety of small molecule ligands was followed by heteronuclear NMR experiments. A general observation is that ligands that act as inhibitors of the enzyme in vitro give rise to specific broadening of a subset of N-H cross peaks, including those of residues around the active site and in the loop that closes the ligand binding pocket. The results are interpreted to imply that in the bound state the dynamic profile of the backbone of the *Pa*DDAH enzyme is altered, probably with ordering of the loop but retaining a component of conformational exchange that gives rise to residual line broadening, even in the saturated state. A similar result is obtained with a ligand covalently bound to Cys249 of the catalytic triad.

"The real voyage of discovery consists not in seeking new landscapes,
but in having new eyes"
Marcel Proust

Acknowledgements

Over the past four years, I have been extremelly fortunate to have met some wonderful individuals who have helped me understand the significance of a recurring theme in my life: it is not *what* you do that is important, but it is *how* you do it that makes any experience worth while. I now have a chance to spread the credits, as well as to take on all the blame.

I would like to start by thanking my supervisor Paul Driscoll for the immense trust and for the freedom given during these years. This combination was probably responsible for turning my experience as a Ph.D. student into a most pleasurable one. I sincerely hope that my years at UCL were just the beginning of a long collaboration.

I owe an immeasurable debt to all the experts for patiently teaching me their skills. My warmest thanks go to Richard for all the excellent comments, suggestions, and criticisms, which have been invaluable not only to this project but to my well-being. I would also like to express my gratitude to Acely for her generosity and kindness. She might not know it but her perception of life is somehow inspiring. I am equally obliged to Mike for helping me settle in the lab and for pointing this project in the right direction. I also wish to acknowledge those who have left the lab, Andrew, Brian, David, Diego, and Mark J. for all their tricks and recipes that have made my life so much easier.

There are many in this department that from behind the scenes have continually encouraged me during these years. I think of Radwan and Anwar for the constant good humour, of Patricia for never failing to cheer up an entire UCL floor with contagious smiles and laughs, and of Juan for all the stimulating and sometimes challenging discussions. As for my most trustworthy collaborator, I could not forget to thank Ramon.

Outside this microcosm, I am ever so fortunate to have Rafael by my side. I could not thank him enough for his love and incredible understanding, but mostly, for his calm in times when our space had been overtaken by the deliria of a writer.

And lastly, I would like to thank my parents and my sister for their seemingly unlimited belief in me. I have grown fond of their various interpretations of my work, which have always helped me to put things into perspective.

Table of Contents

List of Figures

CHAPTER 4

CHAPTER 5

CHAPTER 6

List of Tables

Abbreviations

ABBREVIATION	DEFINITION
1D	One-Dimensional
2D	Two-Dimensional
3D	Three-Dimensional
A.U.	Absorbance Units
ADI	Arginine deiminase
ADMA	N^{η}, N^{η}-Asymmetric dimethylarginine
AGAT	L-Arginine:Glycine amidinotransferase
AMA	Asymmetric methylarginine
ASA	Accessible Surface Area
AST	Arginine succinyltransferase
AT	Amidinotransferase
atRA	Retinoic Acid
COLDER	Colour Developing Agent
CPMG	Carr-Purcell-Meiboom-Gill
CSA	Chemical Shift Anisotropy
CV	Column Volume
DDAH	Dimethylarginine dimethylaminohydrolase
DHFR	Dihydrofolate reductase
FID	Free Induction Decay
H4	Histone 4
Hcy	Homocysteine
HSQC	Heteronuclear Single Quantum Coherence
INEPT	Insensitive Nucleus Enhanced by Polarisation Transfer
IPAT	L-Arginine:inosamine amidinotransferase
kD	kilo Daltons
LB	Luria Bertani
LDL	Low density lipoprotein
L-NMMA	N^{η}-monomethylarginine
MBP	Myelin binding protein
MCS	Multiple cloning site
mRNA	Messenger ribonucleic acid
MSG	Malate synthase G
Ni-NTA	Nickel-charged nitriloacetic acid
$^{n}J_{xy}$	n-bond scalar coupling between nuclei x and y
NMR	Nuclear magnetic resonance
NO	Nitric oxide
NOE	Nuclear Overhauser Effect
NOS	Nitric oxide synthase
OAT	Ornithine aminotransferase
PAD	Peptidyl arginine deiminase
*Pa*DDAH	*Pseudomonas aeruginosa* DDAH

ABBREVIATION	DEFINITION
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PRMT	Protein arginine methyltransferase
R_1	Longitudinal relaxation rate
R_2	Transverse relaxation rate
RF	Radio-Frequency
S/N	Signal-to-noise
S-AdoMet	S-Adenosylmethionine
SD	Standard deviation
SDMA	Symmetric-L-N^{η}, $N^{\eta'}$-dimethylarginine
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
smD1	Spliceosomal protein D1
snRNP	Small nuclear ribonucleoproteins
SW	Spectral Width
TEV	Tobacco Etch Virus
TROSY	Transverse Relaxation-Optimised Spectroscopy
τ_c	Rotational correlation time

AMINO ACID NOMENCLATURE	
Cys 249	Residue type using the standard three letter code (*e.g.* cysteine) and residue number/position
C249S	Mutation of Cys-249 to Ser-249

Chapter 1

DDAH *in vivo* and *in vitro*

Abstract

This chapter describes the role of asymmetric methylarginine (AMAs) residues in the regulation of the production of nitric oxide (NO), an important signalling molecule implicated in the maintenance of the vasculature homeostasis. In mammals, asymmetric methylarginines (AMAs) are known endogenous inhibitors of nitric oxide synthase (NOS), the enzyme responsible for the generation of NO in cells. Over the last decade, the interest in AMAs has increased as their accumulation in plasma and tissues with the subsequent impairment of NO production has been linked to a number of pathologies affecting the cardiovascular system. The metabolism of AMAs is controlled by the enzyme dimethylarginine dimethylaminohydrolase (DDAH). The discovery that DDAH is indirectly involved in the regulation of nitric oxide levels led to the assumption that this enzyme could represent a novel attractive therapeutic target. With the purpose of gaining structural information on this class of enzymes, the crystal structure of a bacterial DDAH homologue was solved. DDAH belongs to a recently identified superfamily of enzymes that modify the guanidino moiety of arginine residues. Several crystal structures of these enzymes are now available and will be described with an emphasis on their structural similarities. Finally, the scope of this thesis will be outlined, as an effort to pursue further structural work on DDAH enzymes with the application of NMR methodology to assist in the discovery of the enzyme's dynamic properties and in the development of new inhibitor molecules.

1.1 The L-arginine:nitric oxide pathway

The role of nitric oxide in the cardiovascular system

The endothelium regulates vascular homeostasis by producing a number of small molecules that act locally in the blood vessel wall and lumen. One of the most crucial mediators in the maintenance of the vascular tone and structure is nitric oxide (NO) (for a review see Vallance and Chan, 2001). The release of NO from the endothelial cells not only maintains constant vasodilation against a number of endogenous vasoconstrictor effectors (Rees *et al.*, 1989), but also prevents the adhesion and aggregation of platelets and white cells to the endothelium (Radomski *et al.*, 1996), and inhibits smooth muscle cell proliferation, contributing to the inhibition of atherogenesis and vessel occlusion (Böger *et al.*, 1996).

In this context NO is formed from the reaction between molecular oxygen and L-arginine catalysed by the endothelial isoform of nitric oxide synthase (eNOS). In addition to the eNOS enzyme, two other NOS isoforms exist, named according to the cell type and conditions under which they were first identified – neuronal NOS (nNOS) and inducible NOS (iNOS). The effects mediated by NO are not confined to the cardiovascular system, as this molecule has also been implicated in neuronal signal transduction and in the immune response (Alderton *et al.*, 2001).

NO has a short life time of less than 4 seconds as it is rapidly oxidized to nitrite and nitrate by oxygenated haemoglobin before being ultimately excreted in urine (Knowles and Moncada, 1992). In the cardiovascular system, the actions of NO are exerted in the vascular smooth muscle cells where it activates guanylate cyclase by binding to the heme moiety of the enzyme and subsequently inducing a rise in concentration of cyclic guanosine-3'-5-monophosphate (cGMP). cGMP is a second messenger known for mediating many of the biological effects of NO (Knowles and Moncada, 1992).

Dysfunction of the endothelial L-arginine:nitric oxide pathway, associated with a reduction in basal or stimulated NO mediated-vasodilation, is thought to be responsible for most cardiovascular diseases, such as coronary heart disease, hypertension, chronic heart failure, vasculopathies associated with diabetes mellitus, hyperhomocysteinemia and artherosclerosis in general (Vallance and Chan, 2001). The mechanisms that account for the reduction of NO production in cardiovascular disease are still unclear, but possibilities include a deficiency of eNOS co-factors due to cardiovascular risk factors such as smoking (Katusic, 2001), the decline in NO bioavailability caused by

its accelerated degradation by reactive oxygen species during oxidative stress (Cai and Harrison, 2000), and overproduction of endogenous inhibitors of NOS in certain disease states, such as in chronic renal failure (Vallance *et al.*, 1992).

The discovery of endogenous asymmetric methylarginines

In 1992, Vallance and co-workers first described the presence of asymmetric methylarginines (AMAs), N^{η},N^{η}-dimethylarginine (ADMA) and N^{η}-methylarginine (L-NMMA), as endogenous inhibitors of NOS isoforms. Although L-NMMA had long been administered as a synthetic NOS inhibitor (Palmer *et al.*, 1988), it was not known to be a naturally occurring compound in plasma and urine. ADMA was subsequently found to exert a similar action to L-NMMA and was present in circulating plasma at concentrations 10 times greater than that of L-NMMA (Vallance *et al.*, 1992).

The existence of ADMA provided an explanation for what was referred as the "L-arginine paradox" (Tsikas *et al.*, 2000a). It was known through a series of experiments led by Cooke and colleagues that the administration of L-arginine in rabbits could reverse the endothelial dysfunction associated with hypercholesterolemia through an increase in NO levels (Cooke *et al.*, 1991). These results were surprising since endothelial cells were found to contain up to 2 mM L-arginine and human plasma concentrations could range from 40-100 μM (Böger and Bode-Böger, 2001). These levels are far higher than the eNOS K_M of ~3 μM for L-arginine (Pollock *et al.*, 1991). The enzyme should have been saturated in endothelial cells and unable to respond to L-arginine supplementation. The presence of potent endogenous inhibitors of NOS in human plasma – ADMA has an IC_{50} of ~4 μM against eNOS *in vitro* (Tsikas *et al.*, 2000b) – suggested why L-arginine could still improve endothelial function by displacing the inhibitory effects of asymmetric methylarginines.

1.2 Synthesis of methylarginines

Characterisation of protein arginine methyltransferases (PRMTs)

The methylation of arginine residues in proteins is a post-translational modification (PTM) known to occur in most eukaryotic cells. The guanidino nitrogens of the arginine within the polypeptide sequence are modified to contain one or two methyl groups (Figure 1.1A) (Gary and Clarke, 1998). The enzymes responsible for this

reaction are protein arginine methyltransferases (PRMTs) and to date eight mammalian PRMTs (PRMT 1-8) have been characterised, although one of them, PRMT 2, identified by homology has no reported activity. Notably, the metabolic cost of methylation is high: 12 ATP molecules are needed in the generation of a methylated product. As a result it seems that the ultimate benefit of this reaction must outweigh its energetic cost for it to be retained through evolution (Gary and Clarke, 1998).

The PRMT enzyme family has been separated into two sub-groups according to the reaction catalysed, the substrate specificity and the tissue distribution. Type I PRMTs catalyse the formation of the asymmetric monomethylarginine L-NMMA and dimethylarginine ADMA residues, whereas type II enzymes generate $N^{\eta},N^{\eta'}$-symmetric dimethylarginine (SDMA), but they can also form mono-methylated L-NMMA (Figure 1.1B). In both cases, the methyl groups added to the arginine residues come from the donor S-adenosylmethionine (S-AdoMet).

The substrate specificities of the two subtypes are different: the only known substrates of type II PRMT are the myelin binding protein (MBP), histones H4, H2A and H3 and spliceosomal proteins smD1, smD3 and SmB, whereas type I PRMTs show broader substrate specificity, methylating multiple RNA-binding proteins, histones H4 and H3, and several other nuclear and heat shock proteins (Bedford and Richard 2005). The type I PRMTs preferentially target proteins with an arginine-glycine-glycine (RGG) tripeptide motif. However, asymetrically methylated arginine residues have also been found in proteins with RXR and RG motifs (R- is the one-letter code for arginine and X- represents any amino acid). Conversely, symmetrically methylated Arg residues are flanked in a GRG consensus sequence, identified in MBP and in spliceosomal small nuclear ribonucleoproteins (snRNP) (Brahms *et al.* 2000), suggesting the importance of residues in position -1 and +2 for determining substrate specificity. Three crystal structures of type I PRMTs are currently available, rat PRMTs 1 and 3 (Zhang and Cheng, 2003, Zhang *et al.*, 2000) and yeast RMT1/Hmt1 (Weiss *et al.*, 2000). These enzymes all share a conserved core structure which contains the binding domain of the S-AdoMet methyl donor and an acidic pocket for the binding of the protein arginine residue (Weiss *et al.*, 2000).

In terms of distribution in the cell, PRMTs are generally ubiquitously expressed, although a closer examination of the mammalian PRMT 1 and 3 expression patterns show that PRMT 1 is predominantly localised to the nucleus, whilst PRMT 3 is

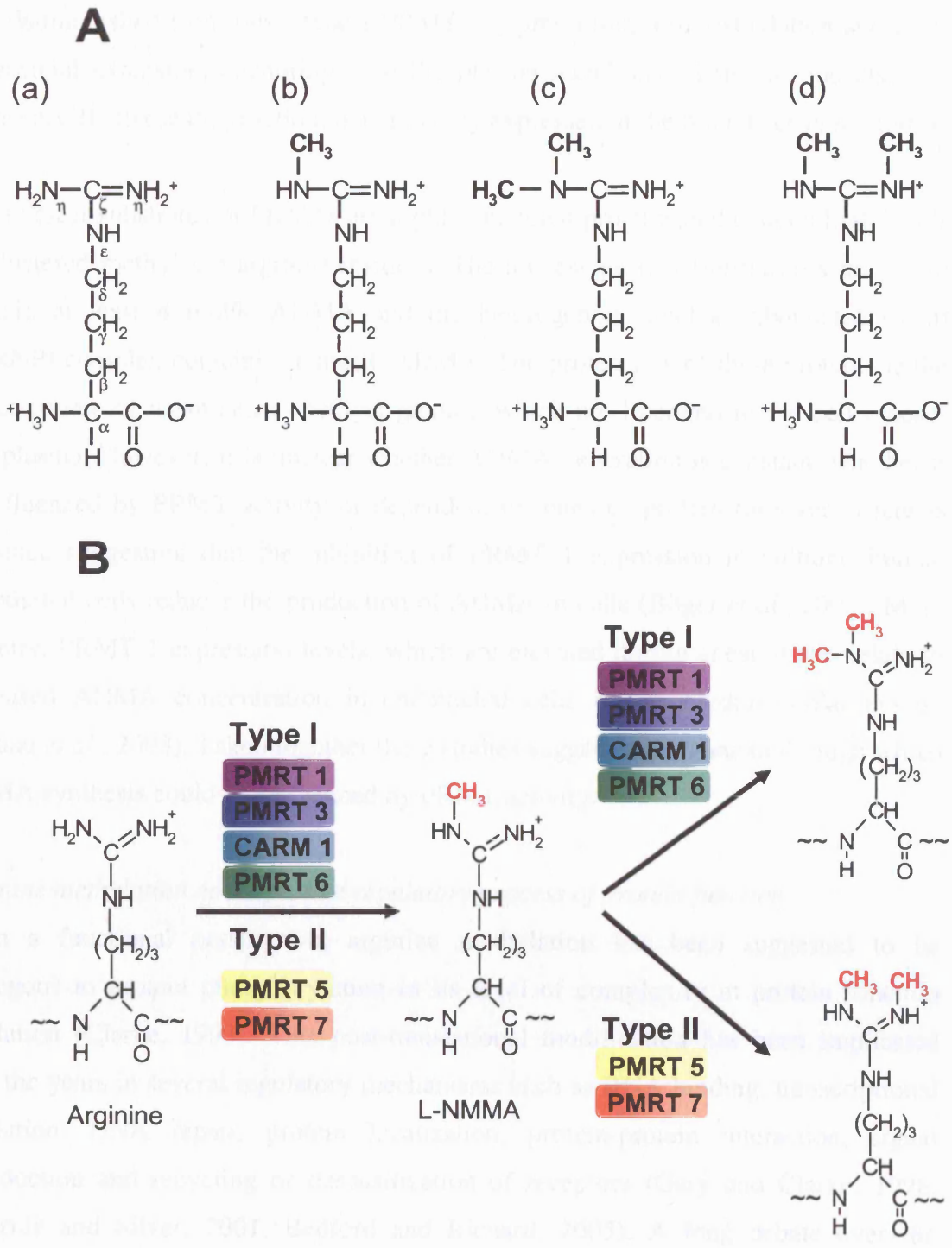

Figure 1.1. (A) Structures of free (a) L-arginine, (b) asymmetric N^{η}-methylarginine (L-NMMA), (c) asymmetric N^{η},N^{η}-dimethylarginine (ADMA) and (d) symmetric $N^{\eta},N^{\eta'}$ dimethylarginine (SDMA). The standard Greek numbering of the side chain carbon and nitrogen atoms is indicated on (a). **(B)** The PMRT family, type I and type II PMRTs generate L-NMMA. Type I alone generates ADMA and the production of SDMA is catalysed by type II PRMTs (adapted from Bedford and Richard, 2005).

predominantly cytosolic (Tang *et al.* 1998). Moreover, the newly characterised PRMT 8 is distinguished from other type I PRMTs by presenting a myristoylation site in its N-terminal extension, anchoring it to the plasma membrane. This enzyme also has more specific tissue distribution, being mainly expressed in the brain (Lee *et al.* 2005).

The nuclear substrates of PRMTs are highly abundant proteins and contain high levels of clustered methylated arginine residues. The nuclear protein fibrillarin is known to contain at least 4 mol% ADMA and the heterogenous nuclear ribonucleoprotein (hnRNP) complex contains ~1 mol% ADMA. The proteolysis of these proteins is the major source of asymmetric methylarginines, which are liberated in the cell cytosol and plasma. However, it is unclear whether ADMA generation is constant, whether it is influenced by PRMT activity or dependent on rates of protein turnover. There is evidence suggesting that the inhibition of PRMT 1 expression in cultured human endothelial cells reduces the production of ADMA in cells (Böger *et al.*, 2000). More recently, PRMT 1 expression levels, which are elevated during shear stress, relate to increased ADMA concentration in endothelial cells of the cardiovascular system (Osanai *et al.*, 2003). Taken together these studies suggest a mechanism through which ADMA synthesis could be influenced by PRMT activity.

Arginine methylation as a dynamic regulatory process of protein function

From a functional perspective, arginine methylation has been suggested to be analogous to protein phosphorylation in its level of complexity in protein function regulation (Clarke, 1993). This post-translational modification has been implicated over the years in several regulatory mechanisms such as RNA binding, transcriptional regulation, DNA repair, protein localization, protein-protein interaction, signal transduction and recycling or desensitization of receptors (Gary and Clarke, 1998, McBride and Silver, 2001, Bedford and Richard, 2005). A long debate over the reversibility of methylation has led to doubts as to whether arginine methylation could act as a dynamic regulatory process of gene expression, since arginine methylation was believed to be an irreversible and stable modification (Bannister *et al.*, 2002). Studies had indicated that arginine and lysine methylation of histones were dynamically regulated (Bannister *et al.*, 2002) but no proof of a demethylation mechanism existed.

Evidence for the reversibility of methylation has only recently been documented. An enzyme named peptidylarginine deiminase 4 (PAD4) was identified as able to convert

all four mono-methylarginine residues of the N-terminal domain of histone 3 (H3) to citrulline (Wang *et al.*, 2004, Cuthbert *et al.*, 2004). This discovery is an important development in the understanding of arginine methylation and demethylation processes, in which the PRMT/PAD pair may act in conjunction, similar to the manner in which other cell signalling enzymes operate, such as kinases and phosphatases (Bannister and Kouzarides, 2005, Denman, 2005).

1.3 Metabolism of asymmetric methylarginines

Methylarginine clearance

Renal excretion was the first mechanism identified for methylarginine clearance. The relatively high quantities of ADMA and SDMA found in urine (Kakimoto and Akazawa, 1970) led to the assumption that ADMA and SDMA levels were good indicators of protein turnover. However, further investigations showed that SDMA levels in urine were in fact thirty times higher than that of ADMA and L-NMMA (McDermott, 1976), suggesting a possible secondary route for the catabolism of asymmetric methylarginines. These results prompted Ogawa and co-workers to isolate and characterise from rat kidney a new enzyme, N^{η},N^{η}-dimethylarginine dimethylaminohydrolase (DDAH), capable of converting free ADMA and L-NMMA residues but not SDMA residues to L-citrulline and methylamines, with K_M values of 180 μM and 360 μM respectively (Ogawa *et al.,* 1989) (Figure 1.2). This enzyme is expressed in kidneys, pancreas, liver and brain (Kimoto *et al.*, 1993) and its expression pattern was confirmed to correlate with that of endothelial NOS and neuronal isoforms in rats. The cloning of the rat DDAH enzyme was followed by the purification and identification of a human DDAH homologue (hDDAH), with 95% amino acid sequence identity (Kimoto *et al.*, 1998). However, the observations that human DDAH activity did not always correlate with the DDAH expression levels led to the discovery of a second novel hDDAH isoform (Leiper *et al.*, 1999). DDAH II has 62 % sequence similarity to the first identified human DDAH sequence, now termed DDAH I (Figure 1.3). The K_M of DDAH II for L-NMMA is 510 μM which is comparable to that of DDAH I (360 μM). The two enzymes have different tissue expression patterns accounting for the initial discrepancies observed between activity and expression levels. DDAH I is highly expressed in brain and kidney, whereas DDAH II is more ubiquitously expressed but mostly found in heart, placenta and kidney. With respect to

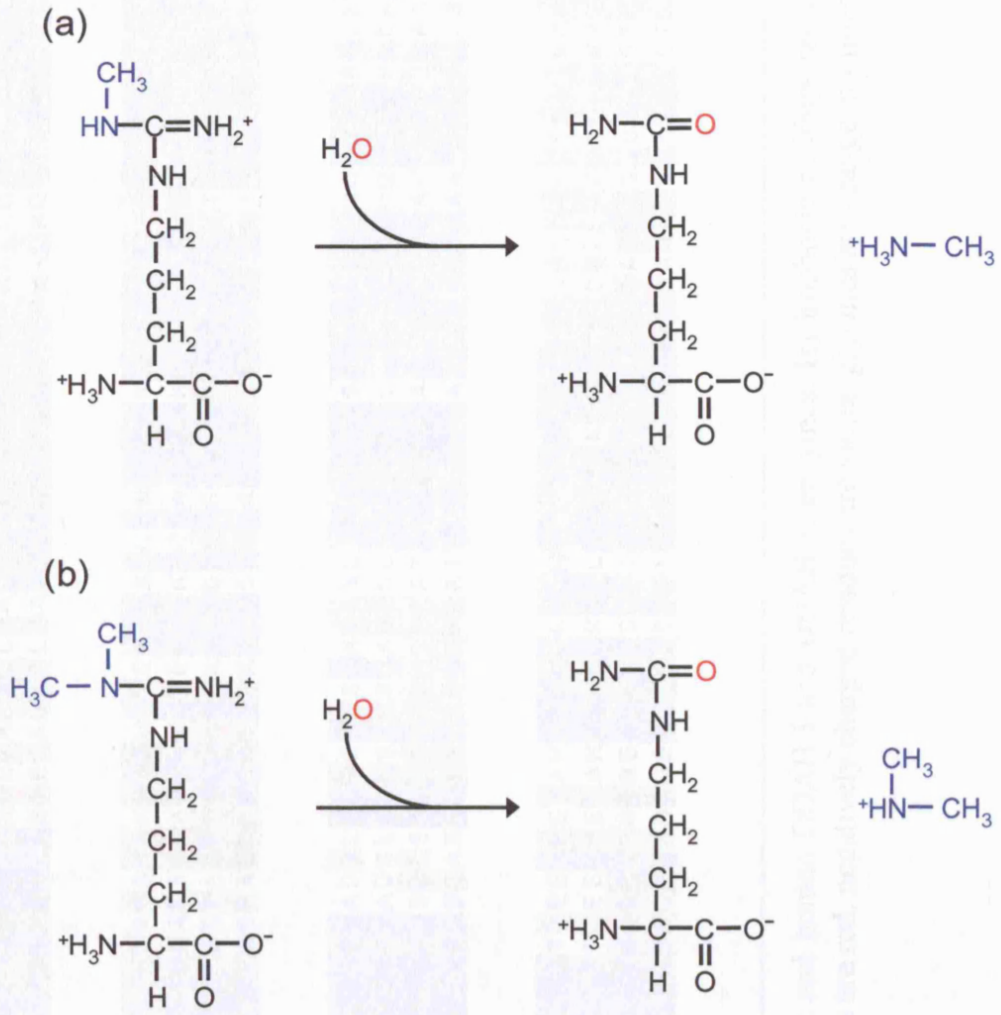

Figure 1.2. Enzymatic reactions catalysed by dimethylarginine dimethylaminohydrolase (DDAH) with (a) monomethylarginine (L-NMMA) and (b) asymmetric dimethylarginine (ADMA) as substrates, L-citrulline and mono- and dimethylamines as products.

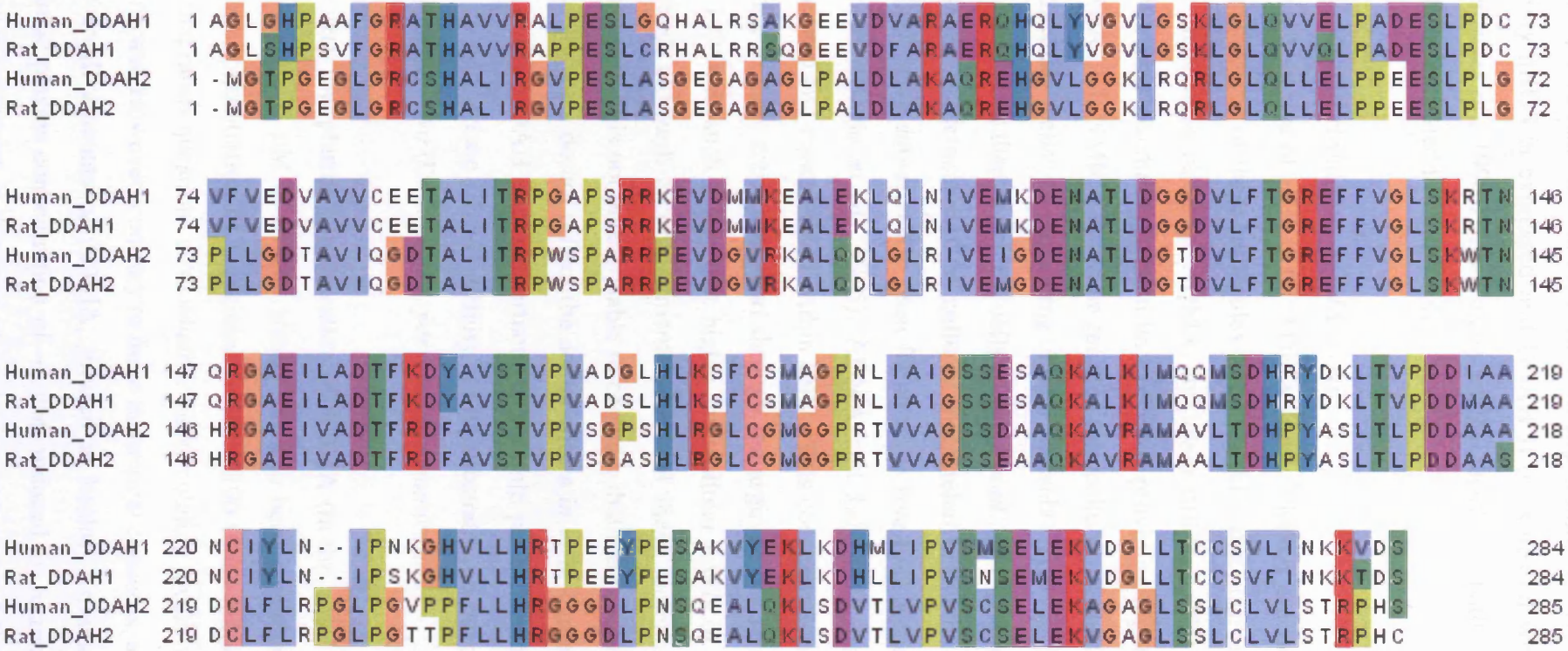

Figure 1.3. Sequence based alignment of rat and human DDAH I and DDAH II enzymes. Hydrophobic residues are coloured blue, polar residues are green, positively charged residues are red, negatively charged residues are violet, glycines are orange, prolines are yellow.

their co-localisation with NOS isoform expression, DDAH I is found in tissues where mainly nNOS is expressed and DDAH II is expressed where eNOS is predominant (Leiper *et al.*, 1999). A summary of the metabolic pathways involving DDAH and NOS is presented in Figure 1.4.

Cellular and circulating ADMA levels

The K_M values of DDAH for ADMA and L-NMMA are rather high compared to the overall levels of these molecules found in cells and plasma, which are usually in the low micromolar range for ADMA, and lower still for L-NMMA (Leiper and Vallance, 1999). This discrepancy could be explained providing higher concentrations of either ADMA or L-NMMA could be reached intracellulary. It has been shown that cultured endothelial cells re-populating blood vessels have higher intracellular ADMA concentrations than in control cells (Azuma *et al.*, 1995). Additionally, as a result of L-NMMA supplementation to cultured endothelial cells, the intracellular levels of L-NMMA have subsequently been found to be five fold higher than in the surrounding medium (Bogle *et al.*, 1995). ADMA and L-NMMA consequently appear to be generated and concentrated within cells. This concentration of methylarginines could be due to their active transport through the arginine Y+ transporter system (Bogle *et al.*, 1995), in such a way that higher intracellular levels of methylarginines could be reached, although a compartmentalisation of these molecules has not been proven. Conversely, it is not unreasonable to consider that the *in vitro* estimations of the K_M of these enzymes do not reflect the actual values *in vivo*. Undergoing research is aiming to identify DDAH binding partners or possible phosphorylation and acetylation sites that could increase DDAH affinity for its substrate and hence lower the K_Ms of these enzymes *in vivo* (Dr J. Leiper, *personal communication*).

The very low plasma concentration of ADMA (in the order of 0.5-1.2 μM in healthy subjects and 3 μM in disease states) has also been a matter of interest. The relative plasma concentration of L-arginine is 60 to 100 times higher than that of ADMA, a fact that raises questions as to whether the circulating levels of ADMA are biologically significant. However, contrary to these theoretical concerns, a recent study showed that very small amounts of ADMA, infused in healthy volunteers in order to reach a systemic plasma concentration of ~2 μM, induced an increase in blood pressure and vascular resistance, a decrease in cardiac output and a fall in cardiac heart rate as a

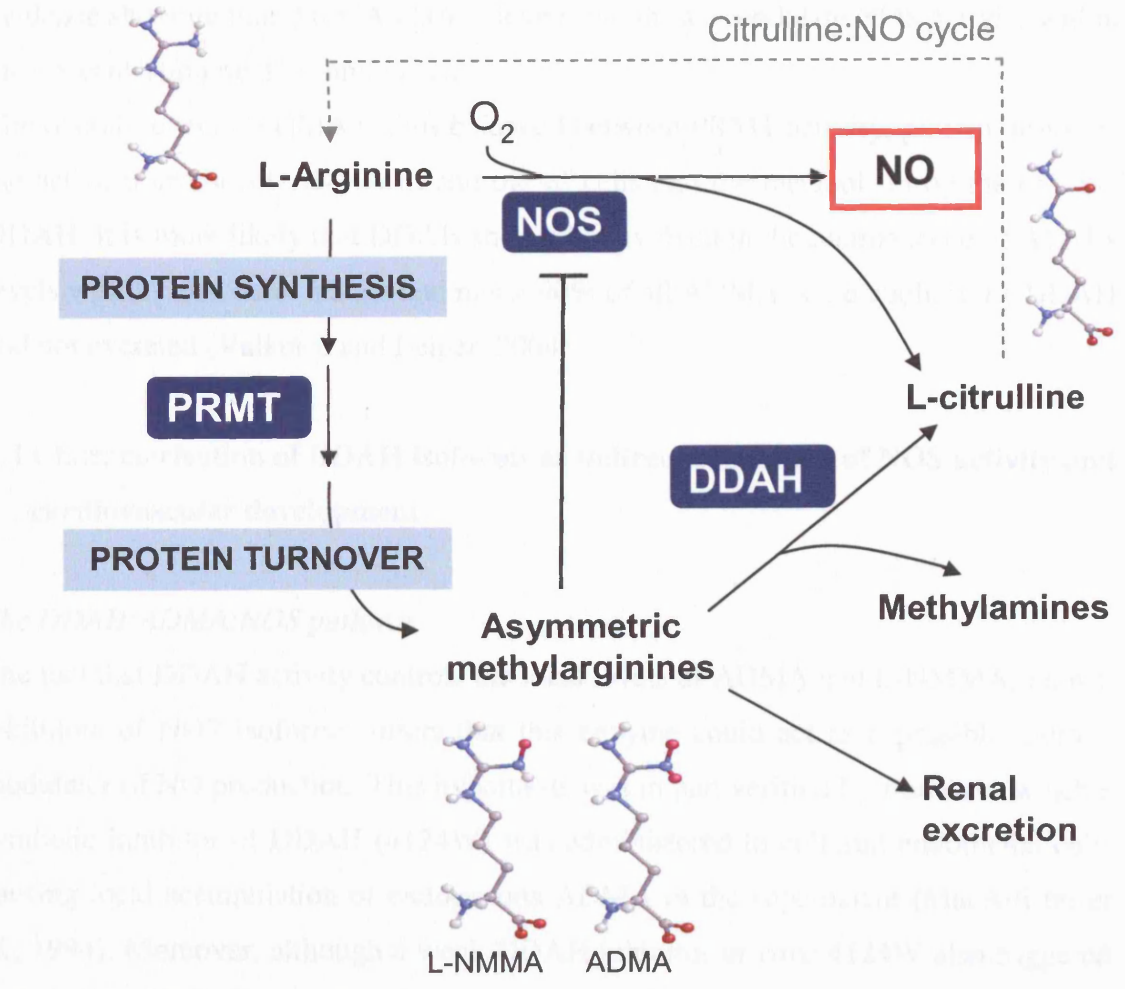

Figure 1.4. Regulation of nitric oxide (NO) and asymmetric methylarginine metabolism. NO is formed by the nitric oxide synthase (NOS) catalysed reaction between molecular oxygen and free L-arginine. Free arginines are incorporated into proteins and can be methylated by protein arginine methyltransferases (PRMT). After protein turnover, endogenous methylarginines are liberated into the cytoplasm, but only asymmetric methylarginines (ADMA and L-NMMA) are known inhibitors of NOS. Dimethylarginine dimethylaminohydrolase (DDAH) converts them into L-citrulline and monomethyl and dimethylamines. Alternatively, they can also follow the renal excretion pathway. L-citrulline is then converted back into L-arginine by urea cycle enzymes (adapted from Leiper *et al.* 1999). The atoms shown are colour coded as follows: carbon – grey, nitrogen – blue, hydrogen – white, oxygen –red and methyl groups are shown in magenta.

direct effect on NOS systems (Achan *et al.*, 2003). These results are conclusive evidence showing that ADMA plasma levels can in fact modulate NOS activity within the concentration range found *in vivo.*

The overall level of ADMA is thus balanced between PRMT activity, protein turnover, the active transport of ADMA in and out of cells and the metabolism by the enzyme DDAH. It is more likely that DDAH should be involved in the maintenance of ADMA levels within cells, as in rabbits and rats > 90% of all ADMA is metabolised by DDAH and not excreted (Vallance and Leiper, 2004).

1.4 Characterisation of DDAH isoforms as indirect regulators of NOS activity and cardiovascular development

The DDAH:ADMA:NOS pathway

The fact that DDAH activity controls the local levels of ADMA and L-NMMA, known inhibitors of NOS isoforms, infers that this enzyme could act as a possible indirect modulator of NO production. This hypothesis was in part verified by studies in which a synthetic inhibitor of DDAH (4124W) was administered to cultured endothelial cells causing local accumulation of endogenous ADMA in the supernatant (MacAllister *et al.*, 1994). Moreover, although a weak DDAH inhibitor *in vitro* 4124W also triggered endothelium-dependent contraction of a rat aortic ring, an effect that was reversed by supplementation of L-arginine. These preliminary findings were further confirmed by studies in transgenic mice where DDAH I was over expressed (Dayoub *et al.*, 2003). The conclusions of this work suggested that the excess of DDAH activity was implicated in the upregulation of NOS activity and NO production, owing to a decrease of ADMA plasma levels, which were reduced by half when compared to controls. DDAH transgenic mice have lower blood pressure and physiological responses consistent with an increase in NO generation. Hence, DDAH is responsible for maintaining the levels of ADMA below the inhibitory concentrations of NOS and modest changes in ADMA levels induced by DDAH activity can have significant effects in NO synthesis and cardiovascular pathology.

Although the human DDAH enzyme expression is not exclusively confined to cells and tissues expressing NOS, these results suggest the existence of a mechanism through which human NOS isoforms could be potentially regulated in a tissue-specific

manner by the modulation of asymmetric methylarginine concentrations. In fact, in view of the systemic physiological effects of NO in mammals, a major pharmaceutical objective is to obtain isoform-specific NOS inhibitors. High levels of NO are often reported in septic shock, neurodegenerative diseases and certain inflammatory conditions. In some instances the blockade of overproduction of NO, generally associated with the expression of inducible NOS (iNOS), is of pharmacological interest (Hobbs *et al.*, 1999). A potential alternative approach to inhibit excess NO production is to target the activity of DDAH. Although the inhibition of one DDAH isoenzyme will not inhibit a specific NOS isoform, it might limit NO synthesis in a tissue-specific manner (Vallance and Leiper, 2002). To this end, several selective substrate-based inhibitors of the mammalian DDAH enzymes have been synthesised which did not inhibit NOS (Rossiter *et al.*, 2005). Two of the designed compounds have an *in vitro* IC_{50} value in the 20 μM range which is a significant advance compared to the millimolar inhibitor 4124W. These molecules are the first reversible DDAH inhibitors described to date and their use will be invaluable in the study of the relevant biological pathways involved.

The activity of DDAH I has been shown to be extremely sensitive to oxidative stress. Ito *et al.* showed that challenging cultured endothelial cells with oxidized small density lipoproteins (oxLDL) and tumor necrosis factor-α (TNF-α), known to decrease NO production in atherosclerosis, significantly decreased DDAH activity and hence induced the accumulation of ADMA in the cells (Ito *et al.*,1999). The impairment of DDAH I activity is not coupled to a decrease in its expression level but rather to direct inhibition of the enzyme's catalytic activity. The connection between homocysteine and ADMA is also relevant. Homocysteine (Hcy) counters the endothelial-dependent vasodilation of the brachial artery through a mechanism that is not fully understood. Hcy was shown to inhibit DDAH activity via a redox mediated mechanism (Stühlinger *et al.*, 2003) and this reaction caused an increase in the levels of ADMA in cells. The authors concluded that this effect could be a possible explanation for Hcy actions *in vivo*. Consistent with the belief that DDAH activity can be regulated through oxidative stress, it was demonstrated that the DDAH I enzyme was inactivated by S-nitrosylation (Leiper *et al.*, 2002). The nitroso group is thought to modify covalently a reactive cysteine residue in the active site. Although speculative, these results suggest the

potential existence of an elegant feedback mechanism through which high NO levels up-regulate the levels of ADMA.

Additional modulators of DDAH II activity

Building on previous analysis of DDAH tissue distribution, a later more comprehensive study was performed in order to estimate effectively DDAH I and DDAH II mRNA levels in 43 adult and 7 foetal human tissues. The quantification of the mRNA levels revealed that DDAH II is highly expressed in foetal tissue (Tran *et al.*, 2000) and it is possible that DDAH II has an important role in cell differentiation and cell cycle control. Achan and co-workers studied the effects of retinoic acid (atRA) on endothelial cell lines. AtRA is an important molecule in the embryonic development of the cardiovascular system and is believed to convey beneficial effects in models of cardiovascular disease. Their data revealed that the implicated mechanism through which atRA could serve as a potential therapeutic in atherosclerosis and angiogenesis involved the up-regulation of DDAH II expression and not of DDAH I, which ultimately produces the desired increase in the levels of NO in endothelial cells (Achan *et al.*, 2002). Further support for DDAH II's role in cellular differentiation showed that this enzyme is transcriptionally regulated by growth- and developmental-factors in processes in which NO is involved (Smith *et al.*, 2003, Jones *et al.*, 2003).

Another interesting point to note with respect to hDDAH isoform tissue distribution is the apparent lack of DDAH I in immune cells and tissues, such as leukocytes, lymph node and bone marrow, where DDAH II is detected. Tran *et al.* observed that DDAH II localisation on chromosome 6p21.3 corresponds to a region containing several genes involved in the immune and inflammatory responses and responsible for the onset of autoimmune diseases, like arthritis or diabetes (Tran *et al.*, 2000). These conditions are linked to an increased activity of inducible NOS in immune cells, where DDAH II gene expression could be implicated in the regulation of methylarginine levels and subsequently NOS expression. The DDAH II gene has therefore been called a "housekeeping" gene or a "disease-susceptible" gene (Jones *et al.*, 2003).

The modulation of DDAH activity appears to be subject to complex regulatory processes. The enzyme expression levels can be manipulated with drugs such as estrogens (Holden *et al.*, 2003), as well as with retinoic acid. The utility of these agents

as tools to probe the therapeutic potential of the regulation of DDAH activity is becoming evident and will likely provide further insight into the biology of ADMA.

1.5 ADMA as a novel marker for endothelial dysfunction and cardiovascular risk

Pathophysiology

High plasma levels of ADMA have been described in a number of diseases associated with the disruption of cardiovascular function, including hypercholestererolaemia (Böger *et al.*, 1998), atherosclerosis (Böger *et al.*, 1997, Miyazaki *et al.*, 1999), hypertension (Surdacki *et al.*, 1999) and chronic heart failure (Usui *et al.*, 1998). Hypercholestererolaemia is a dominantly inherited genetic condition that results in markedly elevated LDL (low-density lipoprotein) cholesterol levels beginning at birth and resulting in premature atherosclerosis of the coronary arteries. The studies conducted by Böger and co-workers showed that high cholesterol levels increased ADMA concentration, although the mechanism responsible for this association remains unknown. The changes in ADMA levels reduced NO production in hypercholestererolaemic subjects and in atherosclerotic patients, substantiated by the reduced urinary nitrate excretion and impaired endothelium-dependent NO vasodilation. The observation that ADMA levels also increase early in the development of atherosclerosis suggests that ADMA is not only a marker of but a mediator of endothelial dysfunction. Moreover, data favouring a role of ADMA as a cardiovascular risk marker is emerging. So far three prospective clinical studies have been carried out with a focus on evaluating ADMA concentration in patients as well as a variety of other cardiovascular risk factors. The main conclusion drawn from these follow-up cases is the strong correlation between elevated ADMA levels and the increased risk of the occurrence of a cardiovascular event (Böger, 2004).

However the highest concentrations of ADMA are found in patients with chronic renal failure. In such individuals, plasma ADMA and SDMA accumulate as they can no longer be excreted. In fact ADMA is present in lower amounts than SDMA, probably due to DDAH activity in the kidneys. The rise in concentration of ADMA was linked to an overall decrease in NOS activity which ultimately led to endothelial dysfunction and unwanted cardiovascular effects (Zocalli *et al.*, 2001). In this study ADMA levels, together with patient age, were the two predictors for overall mortality and cardiovascular events in subjects with renal impairment. ADMA is not only a novel

marker for cardiovascular risk but also emerges as an uremic toxin: it accumulates as the kidney fails, it is a guanidine compound, a product of protein metabolism and it disrupts several biological functions through its effects on NOS function (Kielstein and Zoccali, 2005)

Targets for treatment

One of the methods to counter the actions of ADMA in cardiovascular diseases is L-arginine supplementation. L-arginine has been shown to reverse the effects of ADMA by restoring NOS activity and improving endothelium-dependent vasodilation in hypercholesterolaemic individuals and in patients with congestive heart failure (Böger and Ron, 2005). Nutritional supplementation of L-arginine is sufficient in these cases to restore the physiological ratios of L-arginine to ADMA, as it was shown that L-arginine administration in healthy subjects had no effect on vasodilation. However the scope of these studies is limited and the real potential of L-arginine as a drug remains to be further investigated. Another alternative route to restore ADMA levels is via the up regulation of DDAH activity. As previously shown, estrogens and retinoic acid can increase DDAH expression in experimental models (Holden *et al.*, 2003, Achan *et al.*, 2002). Nevertheless, it is early days before the modulation of DDAH activity proves to be a truly effective therapeutic strategy. The utility of such an approach will certainly benefit from the development of specific inhibitors and activators of DDAH and from studies with animal models in which mice knockouts could provide conclusive evidence for the pathophysiological role of DDAH.

1.6 Evolutionary perspective

The DDAH enzymes are not confined to eukaryotic organisms, the discovery of a family of conserved DDAH enzymes in bacteria indicates the ancestral origin of these enzymes (Santa-Maria *et al.*, 1999). Curiously, many if not all bacteria that harbour DDAH genes are known pathogens. In *Mycobacterium tuberculosis* the gene encoding DDAH is flanked by rocD1, rocD2 and rocE operons, which encode for ornithine handling enzymes such as ornithine aminotransferases (OAT) and for an arginine/ornithine transporter. Ornithine is a biosynthetic precursor of arginine residues, and OATs are important for its metabolism. In *Pseudomonas aeruginosa*, the DDAH gene is part of an operon where putative protein sequences involved in arginine

transport and regulation have been found. Since bacteria are not known to synthesise asymmetric methylarginine residues (for example, the existence of PRMTs in prokaryotes has not been established), the benefit that DDAH enzymes could convey to a bacterium are still unclear. DDAH expression could be a response mechanism to metabolise effectively exogenous ADMA, with re-integration of the L-citrulline residues formed into the general L-arginine metabolic pathway. In fact, *P. aeurginosa* is commonly associated with urinary tract infections (Nicolle, 2002) and ADMA levels are known to be higher than those of L-arginine in human urine (Kakimoto and Akazawa, 1970). Another point to consider is that bacterial DDAH could be implicated in the disruption of NO production in the host. One study has shown that *P. aeruginosa* mediates epithelial damage through NO production but the mechanism through which this effect is achieved is unknown (Downling *et al.* 1998).

Interestingly, a recent report has revealed the presence of a NOS-like enzyme in *Bacillus subtilis* fully capable of producing NO (Adak *et al.*, 2002a) and various NOS oxygenase-like domains have been discovered in other organisms (Adak *et al.*, 2002b). The possible co-existence of NOS and DDAH-like enzymes early on in evolution demonstrates that the metabolism of asymmetric methylarginines and the regulation of NOS iso-enzymes have remained of significant importance. Gene structure comparison of human DDAH I and DDAH II confirms that these two sequences originated from a tandem gene duplication event, proposed to have occurred *ca.* 450 million years ago (Tran *et al.*, 2000).

1.7 Structural characterisation of the *Pseudomonas aeruginosa* DDAH

Description of the overall fold of the Pseudomonas aeruginosa DDAH

As part of ongoing efforts to derive structural information for DDAH, the crystal structure of the more tractable bacterial *Pseudomonas aeruginosa* DDAH (*Pa*DDAH) homologue has been solved (Murray-Rust *et al.*, 2001). Since expression levels for recombinant mammalian DDAH never reached suitable levels for either crystallographic or NMR studies, attention was focused on the bacterial homologue. The enzyme is composed of 254 amino acid residues and has 44% and 22% sequence homology to DDAH I and DDAH II respectively. The overall fold of this enzyme forms a propeller comprised of five modules that enclose a buried active site. The modules are arranged in a pseudo five-fold rotational symmetry (Figure 1.5a,b). Each

module contains two anti-parallel β-strands, followed by an α-helix and a third β-strand that makes a parallel arrangement with the other two β-strands. This motif has been described as the ββαβ structural motif (Figure 1.5c-e). From a lateral view point of the enzyme (with respect to the axis of the barrel), the N- and C-terminal tails of each ββαβ module protrude from the bottom end of the protein. The polypeptide sequence starts with the second β-strand β1B of module 1, followed by helix α1 and the third strand β1C (residues 1-67). The polypeptide chain then continues through all the successive modules finally to close the fold with the C-terminal β-strand β1A (residues 252-254) to complete the initial module 1 (Figure 1.5c,d). This arrangement of the chain is often termed "Velcro closure" and is reminiscent of the β-propeller architecture found in a variety of unrelated proteins. More recently, Paoli suggested that proteins with ββαβ structural motifs such as that found in DDAH should be termed β/α–propellers to differentiate them from the canonical β-propellers which are normally composed of motifs with purely antiparallel four-stranded β-sheets (Paoli, 2001).

In terms of oligomeric state, *Pa*DDAH exists as a homodimer in the crystallographic unit cell. Size exclusion chromatography and dynamic light scattering (DLS) experiments confirmed the dimeric state of the enzyme. The protomer interface surface area was estimated at 860 $Å^2$/protomer. It is located on the side of the enzyme, between helices α1 and α2 and forming an extension of the β-sheet in the ββαβ module 1 across the interface. A detailed description of the dimer interface interactions will be given in Chapter 3.

The catalytic triad and the active site of PaDDAH

Several authors have shown that the *Pa*DDAH β/α–propeller fold is conserved in a variety of enzymes that catalyse modifications of the guanidino group of arginine residues (Leiper *et al.*, 1999, Murray-Rust *et al.*, 2001). Structure based-alignments showed that *Pa*DDAH has sufficient sequence homology to aminidinotransferases (AT) from human and bacteria (discussed in detail in Section 1.8) to provide insights into the residues responsible for the reaction mechanism. The catalytic triad of *Pa*DDAH was thereby identified as being composed of residues Cys 249, His 162 and Glu 114. The single mutant proteins C249S, H162F and E114Q *Pa*DDAHs were correctly expressed in a folded form and they all lacked enzymatic activity. The C249S

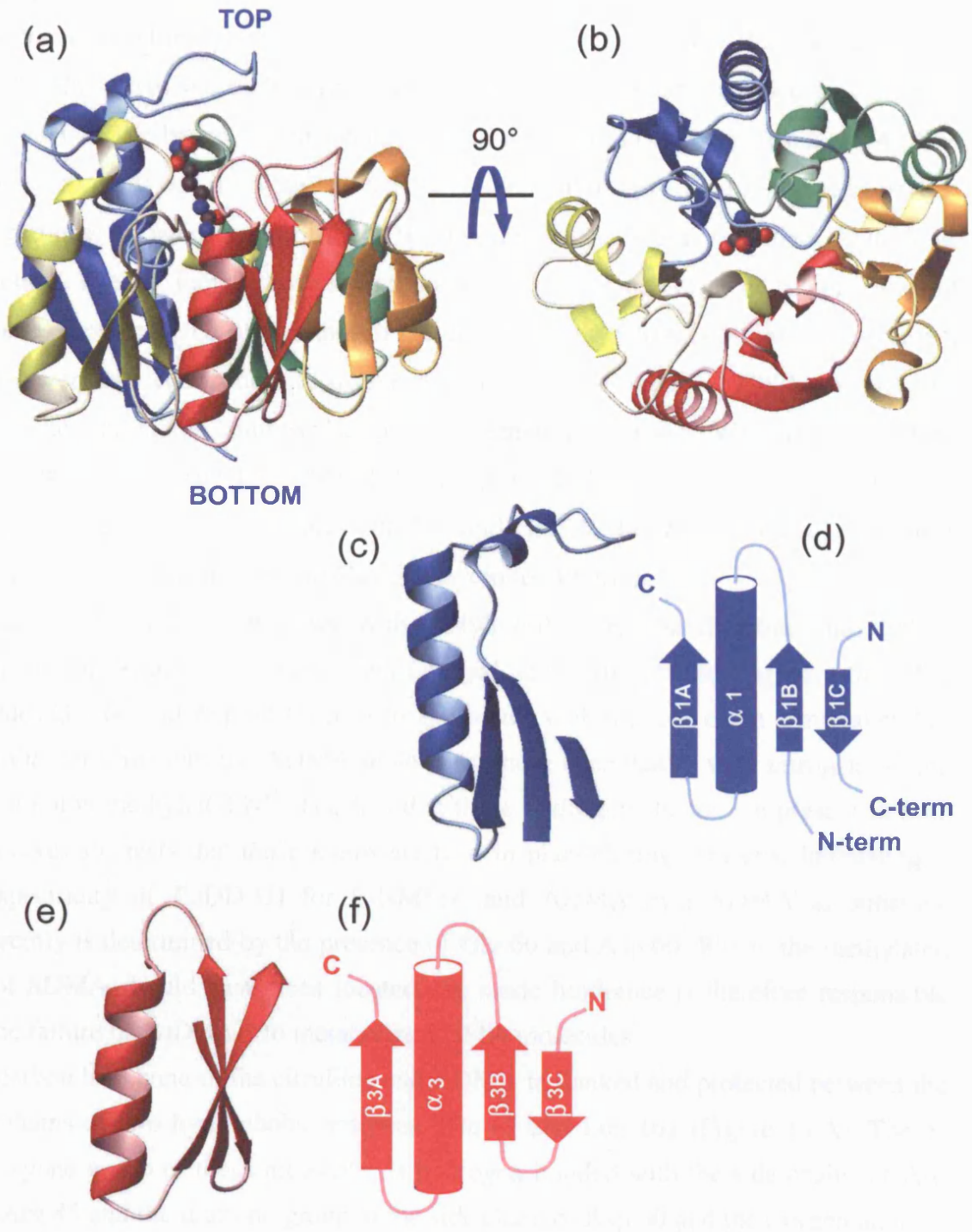

Figure 1.5. 3D structure of the C249S *Pseudomonas aeruginosa* dimethylarginine dimethylaminohydrolase bound to L-citrulline (*Pa*DDAH; PDB code: 1h70). (a) Lateral view of the 3D structure of a single protomer of *Pa*DDAH with ββαβ modules coloured individually (Module 1 is blue: residues 1-67 and 249-254; module 2 is yellow: 68-117; module 3 is red: 118-166; module 4 is orange: 167-205; module 5 is green: 206-248). (b) plane view of a single subunit of *Pa*DDAH coloured as (a). 3D structure (c) and topology (d) of module 1 coloured as (a). 3D structure (e) and topology (f) of module 3. Figures were generated with MOLMOL (Koradi *et al.*, 1996) .

*Pa*DDAH structure was solved with L-citrulline and ADMA in complex, although only the L-citrulline-bound structure has been deposited in the Protein Data Bank (PDB). The active site cleft in the apo-form of the enzyme lies exposed to solvent in the center of the barrel. Upon ligand binding a loop (L1) between residues 14-25 is believed to close down on the active site. A single hydrogen bond is formed between the carbonyl oxygen atom of Leu 18 and the amino group of either ADMA or citrulline. The L1 loop is mobile and disordered in the apo-state of the protein, but becomes more ordered once either of the citrulline or ADMA complexes are formed, with increased electron density in this region (Murray-Rust *et al.*, 2001). In fact, the active site entrance is composed of an arrangement of loops, between residues 54-67 (L2), 78-82 (L3), 107-113 (L4), 130-133 (L5), 157-161 (L6) and 243-251 (L7). Curiously, the active site residues Glu 114 and His 162 lie in 3_{10}-helix segments just adjacent to loops L4 and L6, and Cys 249 lies in the L6 loop.

In each of the ligand complexes with C249S *Pa*DDAH, the citrulline and ADMA molecules are situated in a negatively charged active site pocket (Figure 1.6A). The residues Glu 66 and Asp 60 form hydrogen bonds with the side chain terminal end of both the citrulline and the ADMA molecules, more specifically with nitrogen N^{ε} and with the non-methylated $N^{\eta'}$. The fact that these hydrogens bonds are present in both complexes suggests that these atoms are held in place during catalysis. Interestingly, the specificity of *Pa*DDAH for L-NMMA and ADMA over SDMA as substrate apparently is determined by the presence of Glu 66 and Asp 60. Where the methylated $N^{\eta'}$ of SDMA should have been located, the steric hindrance is therefore responsible for the failure of *Pa*DDAH to metabolize SDMA molecules.

The carbon backbone of the citrulline and ADMA is flanked and protected between the side chains of two hydrophobic residues, Phe 63 and Leu 161 (Figure 1.6A). The α-carboxylate group of these molecules is hydrogen bonded with the side chains of Arg 132, Arg 85 and the α-amino group to the side chain of Asp 60 and the oxygen atom of the carbonyl groups of Arg 132 and Leu 18 (Figure 1.6A).

A catalytic mechanism similar to that of AT has been proposed, in which the nucleophilic attack by the atom S^{γ} of Cys 249 on the C^{ζ} of the asymmetric dimethylarginine substrate is followed by the hydrolysis of the resulting tetrahedral intermediate into citrulline and dimethylamines (Murray-Rust *et al.*, 2001) (Figure 1.6B). The accumulation of a covalent S-alkylthiouronium intermediate attached to the active site cysteine was recently confirmed by acid trapping and mass spectrometric analysis (Stone *et al.*, 2005a).

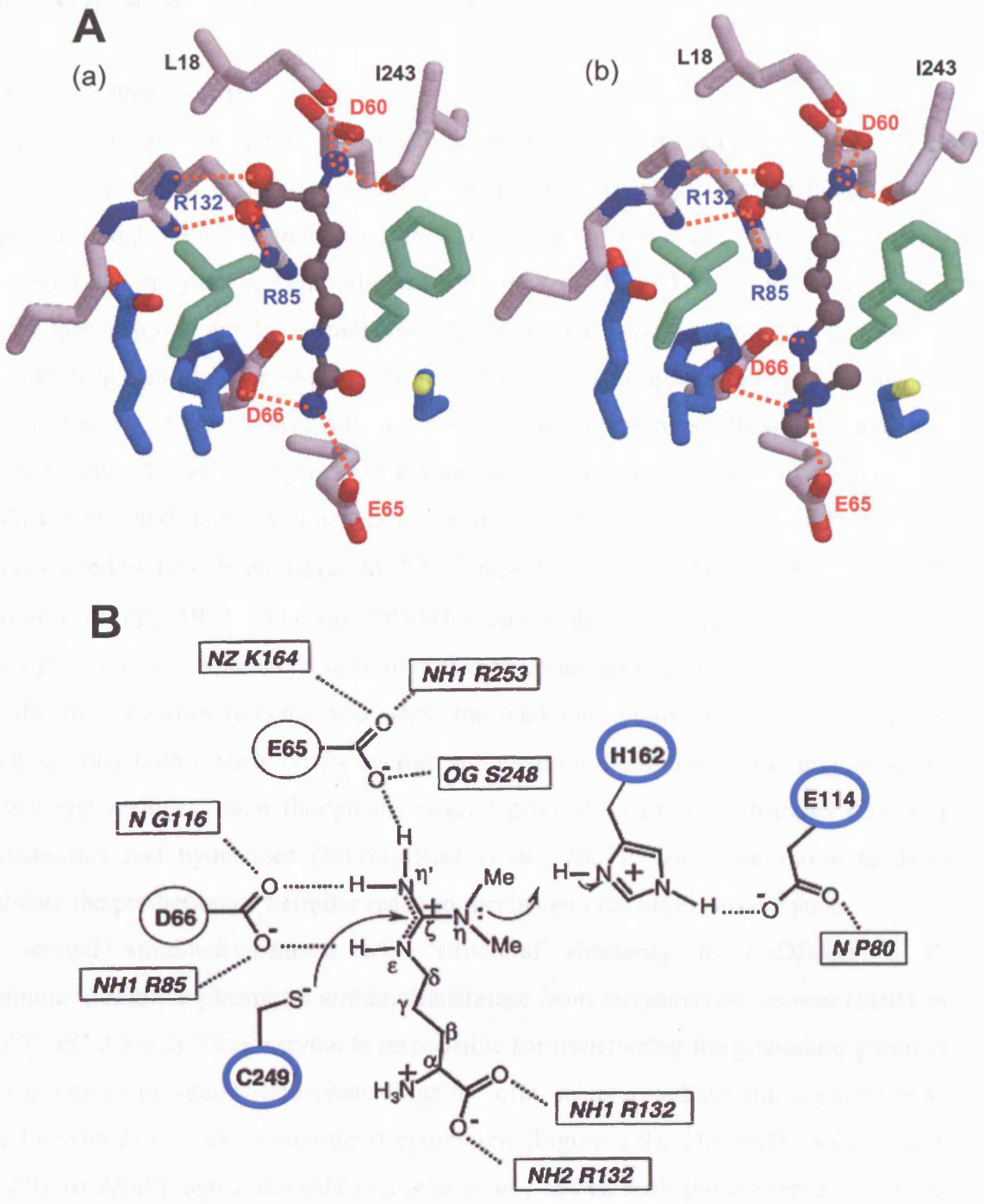

Figure 1.6. (A) 3D view of the active site region of *Pa*DDAH complexed with (a) L-citrulline and (b) ADMA molecules. Active site residues Cys 249, His 162 and Glu 114 are coloured light blue, Phe 63 and Leu 161 are coloured green. All other residues follow CPK colouring. Hydrogen bonds are shown in orange. Note, the O atom of Ser 249 in the *Pa*DDAH pdb file has been coloured yellow to reflect the presence of the S atom of Cys 249. **(B)** Proposed reaction mechanism for the *Pa*DDAH enzyme (Murray-Rust *et al.*, 2001). The C^{ζ} atom of the ADMA molecule undergoes a nucleophylic attack by the thiol group of Cys 249. The reaction intermediate formed is S-alkylthiouronium which is ultimately hydrolysed to form L-citrulline and dimethylamine.

1.8 A superfamily of guanidino-modifying enzymes

The amidinotransferases

Sequence and structure similarity searches for *Pa*DDAH have suggested the existence of a superfamily of arginine-modifying enzymes with strikingly similar β/α–propeller folds, although with low sequence similarity scores (Shirai *et al.*, 2001) (Figure 1.7). Among these enzymes are aminidinotransferases (AT; EC 2.1.4.1), such as the human L-arginine:glycine amidinotransferase (AGAT) that catalyses the transfer of the guanidino group of L-arginine to glycine, producing L-ornithine and guanidinoacetic acid (Figure 1.8). These products are intermediates in the biosynthesis of creatine in humans. AGAT has a catalytic triad similar to DDAH. It is composed of residues Cys 407, His 303 and Asp 254, however the scissile bond of this reaction is different from that cleaved by DDAH enzymes, AGAT cleaves the C^{ζ}-N^{ε} bond of L-arginine residues (Humm *et al.*, 1997) whereas DDAH cleaves the C^{ζ}-N^{η} bond of asymmetric methylarginines. It is interesting to note that the orientation of the active site residues of the two enzymes is conserved when the backbone atoms are superimposed and consequently both scissile bonds overlap and have the same orientation relative to the active site residues, even though the overall orientation of the substrates differs in transferases and hydrolases (Murray-Rust *et al.*, 2001). This observation tends to validate the prediction of a similar reaction mechanism for these two enzymes.

A second amidinotransferase with structural similarity to *Pa*DDAH is L-arginine:inosamine-phosphate amidinotransferase from *Streptomyces griseus* (StrB1 or IPAT; EC 2.1.4.2). This enzyme is responsible for transferring the guanidino group of L-arginine to inosamine-phosphate rings to form an intermediate aminoglycoside in the biosynthesis of the antibiotic streptomycin (Figure 1.9). The StrB1 fold is very similar to AGAT and *Pa*DDAH (Fritsche *et al.*, 1998), with the conservation of the secondary structural elements and position of the active site residues. For StrB1 these are Asp 179, His 227 and Cys 332 (Figure 1.7). However, the AGAT and StrB1 active site regions are shielded by the insertion of three extra α-helices (H1, H2 and H9) in AGAT, two α-helices in StrB1 (H1 and H2) and two extra β-strands at the top end of both molecules, whereas *Pa*DDAH structure only has the L1 loop lying across the active site entrance (Figs 1.8b-1.9b).

Figure 1.7. Structure-based alignment of amino acid sequences from *Pseudomonas aeruginosa* dimethylarginine dimethylaminohydrolase (*Pa*DDAH), *Pseudomonas aeruginosa* arginine deiminase (*Pa*ADI), human peptidylarginine deiminase 4 (HuPAD4), human L-arginine:glycine amidinotransferase (HuAGAT), *Streptomyces griseus* L-arginine:inosamine-phosphate amidinotransferase (SgIPAT) and *Escherichia coli* arginine succinyltransferase (*Ec*ASTR). The colour coding highlights individual ββαβ structural motifs, module 1is blue, module 2 is yellow, module 3 is red, module 4 is orange, module 5 is green. Active site residues are colored light blue and are indicated with the symbol •.

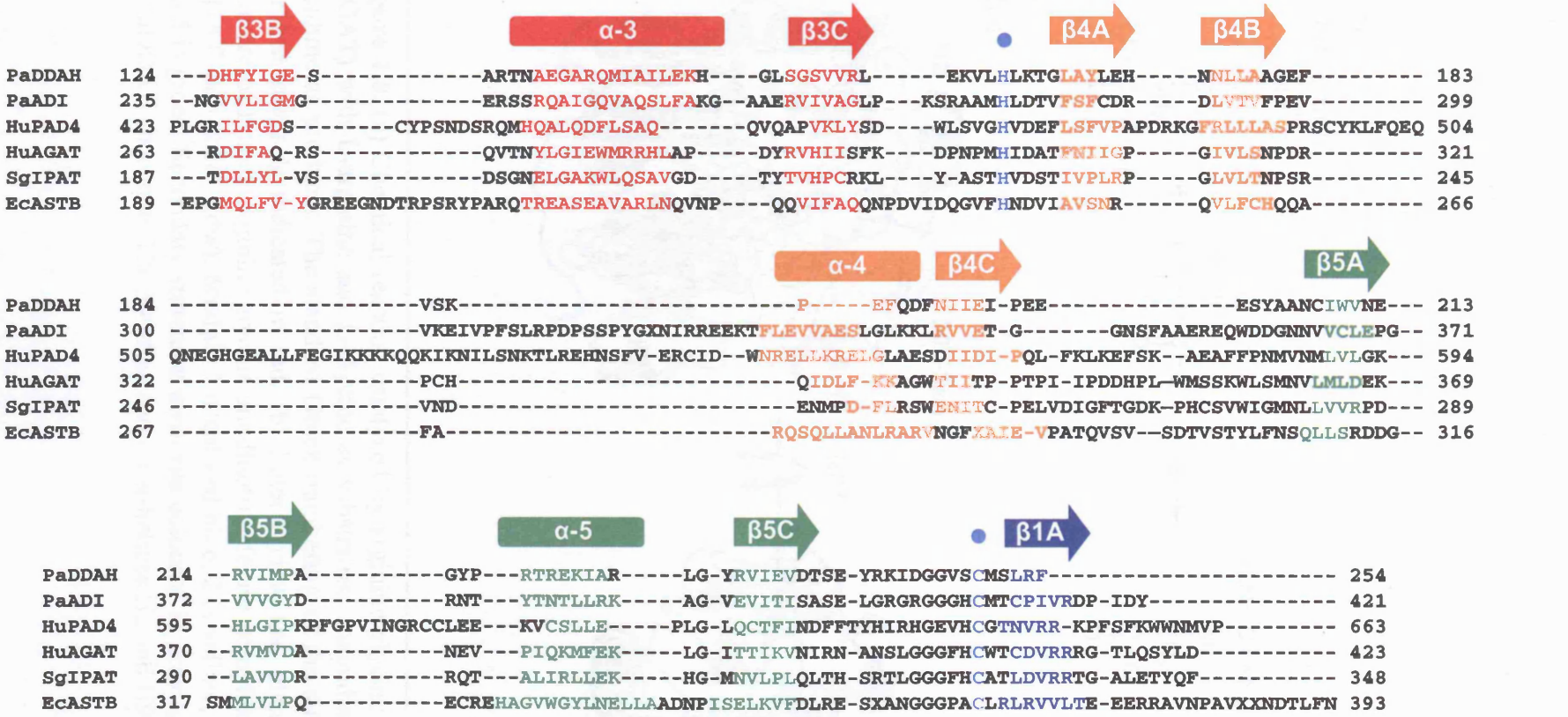

Figure 1.7. *Cont.*

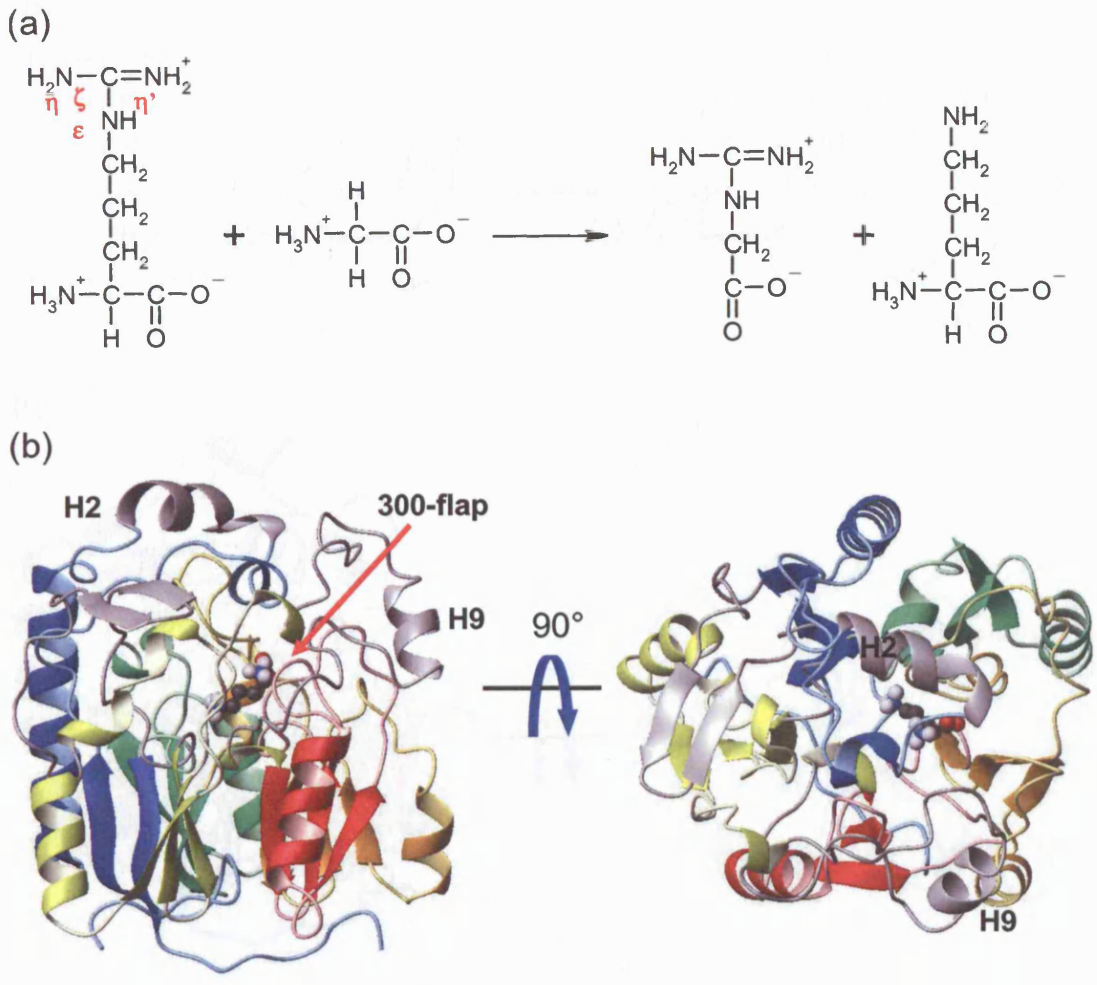

Figure 1.8. (a) Chemical reaction catalysed by arginine:glycine amidino transferase (AGAT) with L-arginine and L-glycine as substrates, guanidinoacetic acid and L-ornithine as products. The standard Greek numbering of the side chain carbon and nitrogen atoms is indicated in red. (b) Lateral view and planee view of the 3D structure of human arginine:glycine amidinotransferase complexed with L-ornithine (AGAT, pdb code: 3jdw). Module 1 is colored blue, 2 is yellow, 3 is red, 4 is orange and 5 is green. Secondary structural elements coloured in grey are insertions relative to *Pa*DDAH sequence. The 300-flap and the α-helices H2 and H9 are shown.

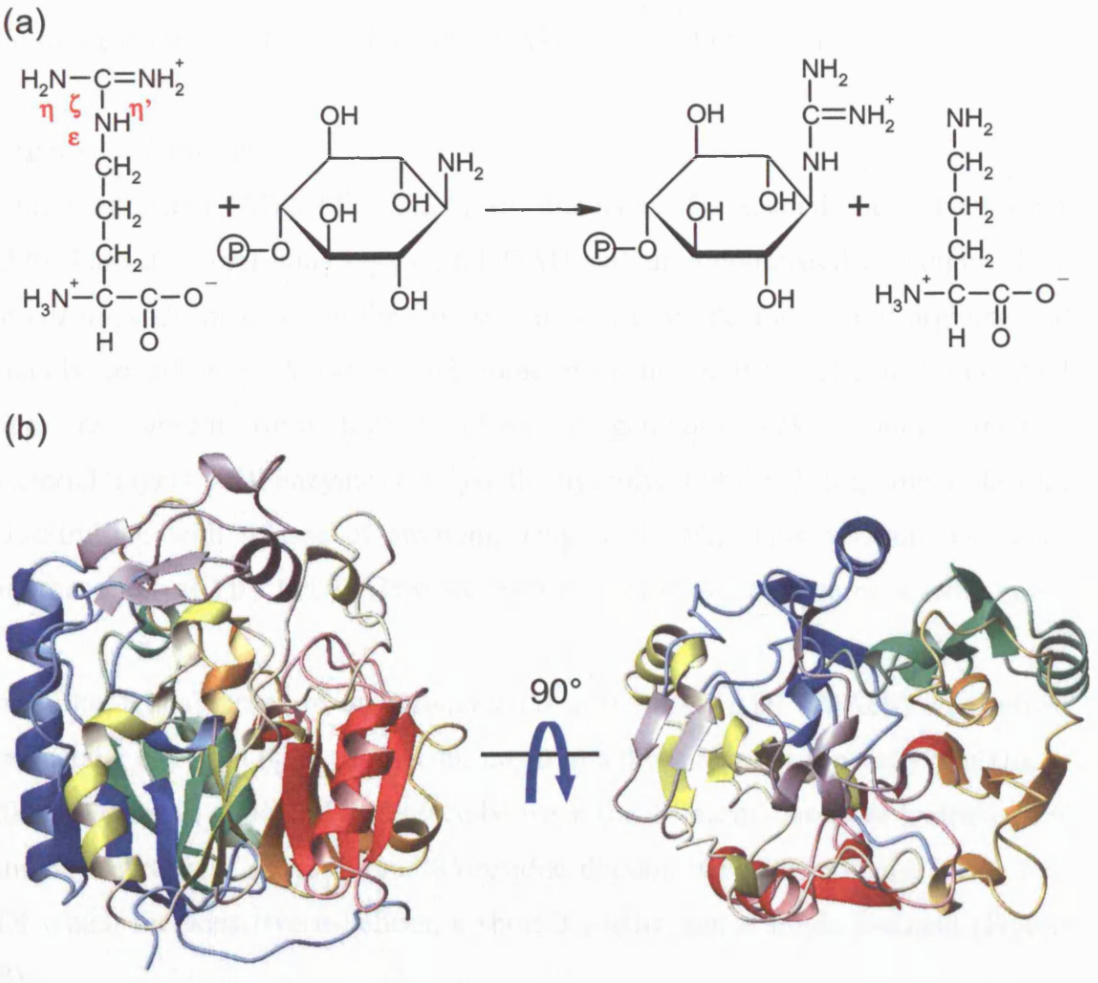

Figure 1.9. (a) Chemical reaction catalysed by arginine:inosamine-phosphate amidinotransferase (AGAT) with L-arginine and *scyllo*-inosamine 4-phosphate as substrates and N-amidino-*scyllo*-inosamine 4-phosphate and L-ornithine as products. (b) Lateral view and plane view of the 3D structure of *Streptomyces griseus* arginine:inosamine-phosphate amidinotransferase (IPAT or StrB1, pdb code: 1bwd). The colour coding is the same as in Fig. 1.8.

The network of surrounding loops is another feature which is not conserved among these three enzymes, but it seems that from crystallographic evidence, the "300-flap" loop between modules 3 and 4 in AGAT (as it is termed by the authors) undergoes a large conformational change upon binding of the substrate (Fritsche *et al.*, 1999), similar to the supposed closing of the *Pa*DDAH L1 loop (Figure 1.8b).

The arginine deiminases

Arginine deiminases (ADI; EC 3.5.3.6) are members of a second class of enzymes found to share structural homology with DDAH and amidinotransferases (Figure 1.7). These enzymes are involved in the first step of anaerobic degradation of arginine and are mainly found in prokaryotes and some parasitic protists. The fact that ADI enzymes are absent from higher eukaryotic genomes makes them attractive antibacterial targets. ADI enzymes catalyse the hydrolysis of free L-arginine molecules into L-citrulline with release of ammonia (Figure 1.10A). This reaction resembles closely that catalysed by *Pa*DDAH since both enzymes target the same scissile bond, C^{ζ}-N^{η}.

Recently the crystal structure of *Pseudomonas aeruginosa* ADI (*Pa*ADI) was solved and revealed a five-fold ββαβ module arranged in a β/α–propeller topology (Galkin, *et al.*, 2004). The main structural difference between the different classes of hydrolases is the unique insertion of an additional 85-residue domain between modules 1 and 2 of *Pa*ADI which includes five α-helices, a short 3_{10}-helix and a single β-strand (Figure 1.10B).

*Pa*DDAH and *Pa*ADI have the same active site residues (the *Pa*ADI catalytic triad is composed of residues Glu 224, His 278 and Cys 406) which closely overlap upon superimposition of the two structures. In addition, the orientations of five other residues in the active site of both enzymes are also in conserved positions relative to the bound substrate (Figure 1.11). This degree of conservation is rather high eventhough the enzymes are capable of discriminating their substrates; DDAH enzymes do not display significant arginine deiminase activity, as the estimated K_M of *Pa*DDAH for L-arginine is in the millimolar range (Stone *et al.*, 2005a). The specificity of DDAH for asymmetric methylarginine residues over L-arginine arises from the positioning of residues Lys 164, Glu 65 and Ser 248 in the *Pa*DDAH active site which are replaced by counterparts Asp 280, Arg 165 and His 405 in *Pa*ADI

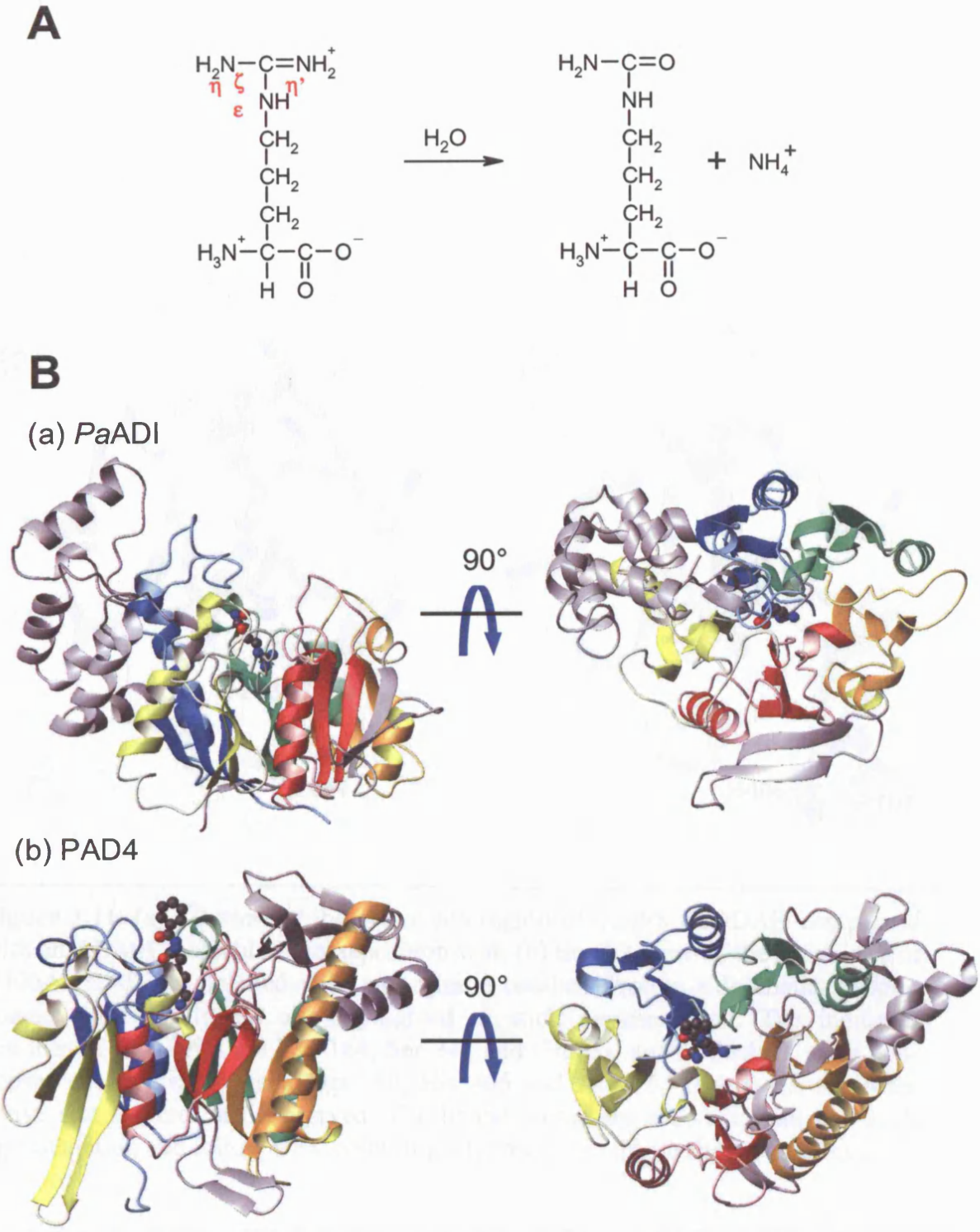

Figure 1.10. (A) Chemical reaction catalysed by arginine deiminase (ADI) with L-arginine as substrate, and L-citrulline and ammonium as products. **(B)** Lateral view and plane view of the 3D structures of (a) *Pseudomonas aeruginosa* arginine deiminase C406A mutant complexed with L-arginine (*Pa*ADI, pdb code: 2a9g), (b) the C-terminal domain of human peptidylarginine deiminase 4 complexed with benzoyl-L-arginine (PAD4, pdb code: 1wda). The colour coding is the same as in Fig. 1.8.

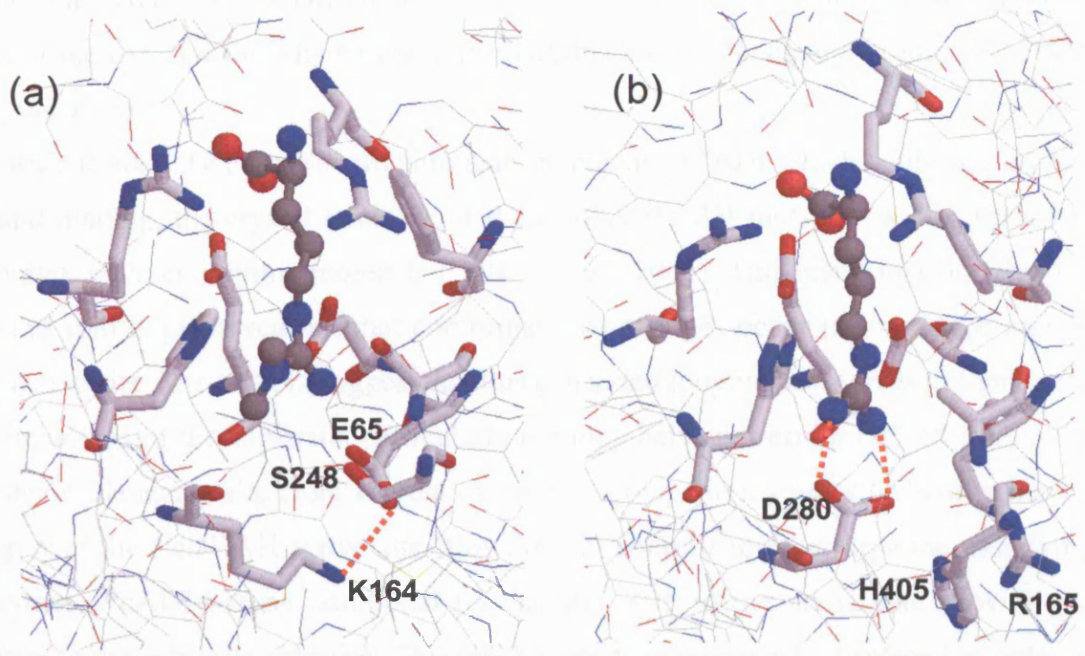

Figure 1.11. (a) 3D view of the active site region of C249S *Pa*DDAH complexed with an ADMA molecule in juxtaposition with (b) the 3D view of the active site of C406A *Pa*ADI complexed with an arginine residue. Protein side chains making contact with the ligand are highlighted in stick representation. The indicated residues in *Pa*DDAH are Lys 164, Ser 248 and Glu 65, and in *Pa*ADI, their non-conserved counterparts are Asp 280, His 405 and Arg 165 (see text). All other active site residues are conserved. The ligand atoms are shown in ball and stick representation and follow CPK colouring. Hydrogen bonds are shown in orange.

(Figure 1.11). If residue Lys 164 in *Pa*DDAH is artificially substituted by Asp 280 as seen in ADI, the carboxyl group of the aspartic acid side chain would clash with the methyl groups of the ADMA molecule. Instead the side chain of Lys 164 in *Pa*DDAH is oriented away from the active site, forming a salt bridge with residue Glu 65, which greatly reduces the steric hindrance and accommodates the methylated arginine in the active site (Galkin *et al.*, 2005). In addition, residue Ser 248 in DDAH, which substitutes His 405 in ADI, allows enough space for the positioning of the aliphatic part of the Lys 164 side chain near to the hydrophobic N^{η},N^{η}-methyl groups of ADMA (Figure 1.11).

In order to identify possible structure rearrangements of ADI which could occur upon ligand binding, the crystal structure of the C406A *Pa*ADI mutant was also solved in complex with an arginine molecule (Galkin *et al.*, 2005). This result revealed that the overall fold is preserved but that conformational changes occur in four loops around the active site. The authors suggest that this concerted loop motion serves to tighten the surroundings of the substrate, particularly the loop between residues 25-48 in *Pa*ADI, to shield the active site from solvent exposure, perhaps in a similar fashion to the L1 loop over the *Pa*DDAH active site. However, in contrast to other arginine-modifying enzymes, *Pa*ADI has an additional residue Arg 401 in the active site providing a barrier to the substrate entrance. This residue needs to be entirely displaced in order to accommodate the arginine substrate. In fact, the end of Arg 401 side chain is displaced by more than 8 Å upon L-arginine binding (Galkin *et al.*, 2005).

The *Pa*ADI reaction mechanism has been investigated in detail with the co-crystallisation of the inactive H278A ADI mutant with the alkylthiouronium reaction intermediate (Galkin *et al.*, 2005). The side chain Cys 406 is responsible for the nucleophilic attack on the carbon C^{ζ} of arginine. A covalent adduct (Cys406-S-alkylthiouronium) is then formed, with the subsequent release of ammonia. The hydrolysis proceeds with a proton transfer to His 278 that enables the production of L-citrulline.

Peptidylarginine deiminases (PAD; EC 3.5.3.15) are yet another family of arginine deiminases, which catalyse the conversion of protein arginine residues to citrulline and ammonia (Figure 1.7). This conversion is termed citrullination and is part of a post-translational modification in vertebrates (Vossenaar *et al.*, 2003). The crystal structure of the human PAD4 enzyme has been solved recently (Arita *et al.*, 2004) and has two

independent domains. The N-terminal domain consists of residues Met 1-Pro 300 and was shown to be composed of two immunoglobulin-like subdomains. The C-terminal domain comprises residues Asn 301-Pro 663 and has structural homology to the other enzymes presented so far with five ββαβ modules also arranged in a β/α–propeller (Figure 1.10B). This enzyme is the first in this class to require the binding of two Ca^{2+} ions for catalysis, and was solved in complex with a benzoyl-L-arginine molecule. It is noteworthy that the major backbone rearrangements of this enzyme appear to occur upon Ca^{2+} binding and not upon substrate binding (Arita *et al.*, 2004). Candidates for the active site residues were deduced from the superimposition of PAD4 and *Pa*ADI 3D structures and were later confirmed to be Asp 350, His 471, Asp 473 and Cys 645. It is interesting to note that the C-terminal β/α–propeller domain has an exposed cavity and no loops are believed to close down on the substrate in the active site, probably due to the fact that the entrance to PAD4 the active site needs to accommodate a polypeptide chain attached to the substrate arginine. By corollary, it is possible that steric effects explain why ADI and AT enzymes can not bind protein arginine residues. Recently, as mentioned in Section 1.2, Wang *et al.* and Cuthbert *et al.* provided evidence that PAD4 can also convert N^{η}-monomethylarginine residues to citrulline, which would imply a similar active site arrangement for both DDAH and PAD4 enzymes (Wang *et al.*, 2004, Cuthbert *et al.*, 2004). However, *in vitro* studies carried out by another group showed that PAD4 poorly catalyses the deimination of protein methylated arginine residues (Kearney *et al.*, 2005). The authors argue that methylated arginines are poor substrates for PAD4 since the active site is too small to accommodate the methyl group. Nevertheless, the apparent contradiction might in part be due to the presence of interacting proteins *in vivo* that could enhance the ability of PAD4 to catalyse the citrullination of methylated arginines. Another argument put forward is the possibility of PAD4 to undergo a post-translational modification that would also enhance PAD4 affinity for these substrates (Kearney *et al.*, 2005). These studies indicate that an additional level of regulation must exist *in vivo* so that the deimination of monomethylated arginine residues remain of physiological relevance as proven in Wang and Cuthbert's independent works.

The arginine succinyltransferase pathway

L-arginine has several routes for catabolism in prokaryotes, among which are the arginase pathway, the arginine deiminase pathway (which was briefly described in the previous section) and the arginine succinyltransferase (AST) pathway. In *Escherichia coli* the AST pathway accounts for 97% of arginine metabolism. The pathway consists of five catalytic steps to convert arginine to glutamate. One such intermediate step is the conversion of *N*-succinylarginine to *N*-succinylornithine catalysed by succinylarginine dihydrolase (AstB, EC 3.5.4.-) (Figs. 1.7 and 1.12a). This enzyme was predicted to be a member of the guanidino-modifying enzyme superfamily through the use of a combination of several computer-based fold and homology recognition tools (Shirai and Mizuguchi, 2003). The crystal structure of *E. coli* AstB solved recently confirmed these findings and revealed the same five ββαβ structural motifs arranged in a propeller and a conserved catalytic triad composed of residues Glu 274, His 248 and Cys 365 (Tocilj *et al.*, 2005, Figure 1.12b). Another feature apparently shared between AstB and DDAH enzymes is the presence of a highly mobile loop, Ala 19 to Gln 34 in the apo-state of AstB, which becomes organized upon substrate and product binding.

AstB was proposed to catalyse two cycles of hydrolysis, removing successively the N^{η} and $N^{\eta'}$ groups from the guanidium moiety of the arginine derivative with release of ammonia, CO_2 and leaving the ornithine side chain as a product (Figure 1.12a). Although AstB cleaves the C^{ζ}-N^{ε} bond just as ATs do, the side chains of AstB residues in the vicinity of the active site share similar environments to ADI active site residues. The question arises as to what keeps AstB from catalysing the conversion of N-succinylarginine to N-succinylcitrulline but instead drives the reaction further through a second hydrolysis step. The only explanation presented by the authors focuses on a single residue substitution in the active site region. In ADI enzymes the residue in question is an aspartic acid (Asp 166 in *Pa*ADI) whereas in AstB dihydrolases the position is always occupied by an asparagine residue which forms only one hydrogen bond to the N^{ε} of the arginyl moiety. Asp 166 in ADI forms instead two hydrogen bonds to the arginine molecule with N^{ε} and N^{η} hampering the second hydrolysis step (Tocilj *et al.*, 2005).

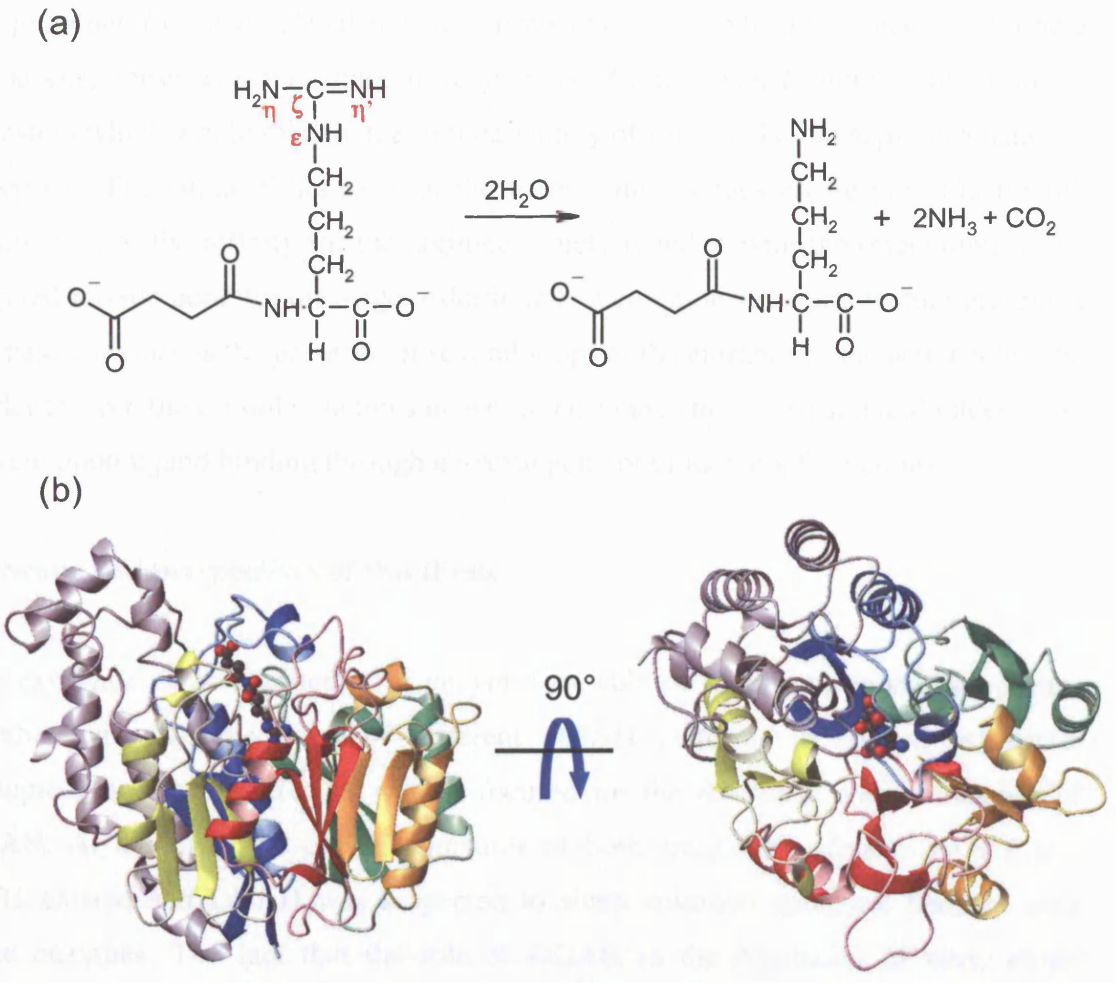

Figure 1.12. (a) Chemical reaction catalysed by succinylarginine dihydrolase (AstB), with L-succinylarginine as substrate, L-succinylornithine, ammonia and carbon dioxide as products. (b) Lateral view and plane view of the 3D structures of *Escherichia coli* succinylarginine dihydrolase C365S mutant bound to succinylarginine (AstB, pdb code: 1yni). The colour coding is the same as in Fig. 1.8.

Concluding remarks

The guanidino group of arginine residues has a high energy content for which nature has devised several pathways of catabolism. The set of enzymes presented here are just a sample of arginine guanidino-modifying enzymes which share a highly conserved β/α-propeller fold and related reaction mechanisms. Despite the diverse amino acid sequences, these enzymes may have evolved from a small number of common ancestors which would explain the limited variety of folds and of catalytic mechanisms described. The subtle differences in the active site residues are responsible for the modulation of the affinity for the arginine moiety together with the orientation of the targeted scissile bond for cleavage. Additionally a common structural feature presented by these enzymes is the presence of several loops at the entrance to the active site. The evidence from the crystal structures is that in most cases the active site is shielded from solvent upon ligand binding through a rearrangement of loop conformations.

1.9 Scope and perspectives of this thesis

The existence of a superfamily of enzymes capable of modifying guanidino groups together with the discovery of different DDAH isoforms in several organisms prompted the idea for a pilot project focused on the structural characterisation of DDAH. At the time, the crystal structures of both amidinotransferases AGAT and SrtB1 existed and DDAH was suspected to share common structural features with these enzymes. The fact that the role of DDAH in the regulation of nitric oxide production was only beginning to emerge and that this enzyme could be employed as an additional strategy in the modulation of NO synthesis only confirmed the need for a superior structural characterization of the enzyme. The efforts were successful in that following a preliminary investigation into the folding status of *Pa*DDAH by ^{1}H NMR, the crystal structure of the bacterial *Pseudomonas aeruginosa* DDAH was solved by colleagues in the Bloomsbury Centre for Structural Biology (Murray-Rust *et al.*, 2001), whilst ongoing work aiming to apply NMR methodology to the study of DDAH was being carried out in our laboratory.

From an NMR perspective, the further study of *Pa*DDAH would have to overcome the large size of the homodimer enzyme (~58 kD), which resides in the upper limits for the useful application of NMR. Signal overlap, increased linewidth, accelerated loss of phase coherence, and low signal-to-noise (S/N) ratio are some of the problems

encountered when dealing with progressively larger proteins by NMR. However, recently developed NMR experiments such as transverse relaxation optimised spectroscopy (TROSY) (Pervushin *et al.*, 1997; Riek *et al.*, 2000; Fernandez and Wider, 2003), coupled with the use of ^{2}H isotope labelling of non-exchangeable protons (Anglister *et al.*, 1993, Gardner *et al.*, 1998) represent a major breakthrough in overcoming the many challenges NMR spectroscopists face when studying large proteins. These two developments have allowed the complete backbone resonance assignments of a 110 kD homooctamer (Salzmann *et al.* 2000) and an 80 kD bacterial malate synthase G (Tugarinov *et al.* 2002), validating the concept that larger molecular weight proteins are tractable by NMR. However, the application of these techniques by a predecessor in our laboratory failed to provide a complete set of resonance assignments for *Pa*DDAH (Plevin, 2003). It was supposed that the intrinsic physico-chemical properties of the homodimer contributed to the apparent absence of a significant number of the expected cross peaks. Our inability to pursue the NMR characterisation of the 58 kD homodimer led us to consider a different direction for our studies.

The crystal structure of *Pa*DDAH revealed that the interface surface area between protomers was small and size exclusion chromatography (SEC) preliminary results showed that *Pa*DDAH exhibited a monomer-dimer equilibrium in solution (Murray-Rust *et al.*, 2001). These findings led us to speculate that the homodimeric nature of the protein could be an artefact of crystallisation and non-essential for the biological activity of the enzyme. In this respect, Chapter 3 will highlight the exploration of the interface surface area of *Pa*DDAH. The identification of key residues in the protomer-protomer contact area allowed for the introduction of point mutations to design a fully active monomeric enzyme, essentially more tractable for heteronuclear NMR studies. Chapter 4 will focus on the successful backbone resonance assignments of the engineered monomeric enzyme. These assignments have helped us to understand the reasons for our inability to obtain unambiguous assignments of the wild-type *Pa*DDAH. In Chapter 5, the intrinsic backbone dynamics of the apo-state of the monomeric *Pa*DDAH will be addressed through ^{15}N relaxation studies. Lastly, Chapter 6 will assess the interaction of *Pa*DDAH with a small library of compounds designed to derive specific inhibitors of mammalian and bacterial DDAH isoforms.

Chapter 2

Material and Methods

Abstract

This chapter describes the experimental protocols and the techniques routinely employed in this work. When further explanations are needed, the reader will be referred to the relevant section of this thesis.

2.1 Chemicals and laboratory equipment

Chemicals used in this work have been purchased from Sigma-Aldrich, BDH *Anala*R or Melford Labs Ltd. All buffers and solutions were prepared with deionised water from a Elga Maxima Ultra pure water purification system.

Absorbance measurements were obtained with a Hitachi U-1800 UV/Vis spectrophotometer. The extinction coefficients for *Pa*DDAH proteins were calculated with the Expasy Server Sequence Analysis Tools (http://www.expasy.ch/tools/) based on the predicted amino acid sequence of the proteins.

Centrifugation steps were performed with a MSE microcentaur (all microfuge tubes), or with an Eppendorf 5810 or Heraeus Labofuge 400R (15 and 50 ml Falcon tubes), or Sorvall RC5B centrifuges (200 ml centrifuge tubes). In the latter case, SS34 and GS3 rotors were used for centrifugation.

SDS-polyacrylamide gel electrophoresis (SDS-PAGE) was performed with BioRad equipment and with the XCell *SureLock*™ system from Invitrogen. All SDS-polyacrylamide gels were prepared as indicated in the Sambrook manual (Sambrook and Russell, 2000) using the BioRad self-assembly gel kit. Alternatively, pre-cast gels (NuPAGE® Novex 4-12% Bis-Tris 1.0 mm, 15 wells) were purchased from Invitrogen.

2.2 General molecular biology protocols

Producing competent cells

E. coli BL21(DE3) competent cells were prepared in house. 100 ml of Luria-Bertani (LB) medium (per liter: 10 g bacterial peptone, 5 g yeast extract, 5 g NaCl) was inoculated with non-competent BL21(DE3)-type cells and grown overnight at 37°C at 200 rpm. The cells were harvested by centrifugation at 2000 rpm for 10 minutes and gently resuspended in 20 ml of pre-chilled sterile 0.1 M $CaCl_2$ solution. The suspension was left on ice for 10 minutes and centrifuged at 1500 rpm for 5 minutes. The cycle of resuspension on ice and centrifugation was repeated four times. The cells were finally resuspended in 4 ml of 0.1 M $CaCl_2$ and 10 % glycerol. 250 µl aliquots of cells were snap frozen in liquid nitrogen and stored at -80°C.

Transformation of competent cells

Transformation of competent cells proceeded as follows: 2 µl of purified plasmid was added to 50 µl aliquots of thawed competent cells and left on ice for 30 minutes. The solution was subjected to a heat shock dependent on the cell type. BL21-type cells were heat shocked for 90 seconds at 42°C and DH5α cells (Invitrogen) were heat shocked for 30 seconds at 37°C. Both cell types were then chilled on ice for 2 minutes before the addition of 250 µl of rich SOC medium (Sambrook and Russell, 2000). The cells were allowed to grow for 30 minutes to one hour at 37°C before being plated. 50 to 100 µl of cell mixture was plated onto a carbenicillin selective 25 ml LB-agar plate and left overnight at 37°C.

Mutation of WT PaDDAH

The cDNA of WT *Pa*DDAH was subjected to site-directed mutagenesis using the Quickchange methodology (Stratagene). Reaction protocols were obtained from the manufacturer's manual. The primers used to generate the mutations used in this thesis are shown in Table 2.1. Oligonucleotides were obtained from MWG-Biotech AG (Ebersberg, Germany). All PCR steps were performed on a MWG Primus thermocycler.

Table 2.1. PCR primers designed for site directed mutagenesis

Mutation	Direction	Primer Sequence (5' to 3')
N36W	Forward	GCCCTGGAGCAGCACTGGGCCTACATCCGCGCC
	Reverse	GGCGCGGATGTAGGCCCAGTGCTGCTCCAGGGC
R40E	Forward	GCACAACGCCTACATCGAGGCCTTGCAGACCTGCG
	Reverse	CGCAGGTCTGCAAGGCCTCGATGTAGGCGTTGTGC
Q43R	Forward	CCTACATCCGCGCCTTGCGGACCTGCGACGTGG
	Reverse	CCACGTCGCAGGTCCGCAAGGCGCGGATGTAGG
R98H	Forward	GGAAACCGTGCAGCACTTCTATCCGGGC
	Reverse	GCCCGGATAGAAGTGCTGCACGGTTTCC

The DNA reaction mixture was transformed into XL21-Blue supercompetent cells (Stratagene). Single colonies, showing successfully transformed cells, were used to inoculate 10 ml LB-carbenicillin medium and incubated overnight at 37°C, 250 rpm. After pelleting the cells by centrifugation (2000 rpm, 10 minutes) plasmid DNA was extracted using QIAprep spin miniprep kits (Qiagen) as instructed in the product manual. DNA sequence analysis was performed by MWG-Biotech sequencing services.

2.3 Expression and purification of *Pa*DDAH proteins

Expression vector for PaDDAH proteins

The cDNA insert encoding the enzyme dimethylarginine dimethylaminohydrolase from *Pseudomonas aeruginosa* has been cloned into the pPROEX Hta plasmid (Invitrogen) between the 5' *Eco*RI and the 3' *Hin*dIII restriction sites. *Pa*DDAH is expressed with an N-terminal polyhistidine affinity tag and a tobacco etch virus (TEV) protease cleavage site to allow the removal of the affinity tag (Figure 2.1). The pPROEX vector contains the β-lactamase gene to confer resistance to ampicillin and carbenicillin.

Large scale expression of unlabelled PaDDAH

A single colony selected from a plate of freshly transformed BL21(DE3) cells with the pPROEX Hta plasmid carrying the cDNA insert of *Pa*DDAH was used to inoculate an overnight 5 ml LB culture containing 100 μg/mL carbenicillin. The cells were incubated at 37°C and 250 rpm. The small culture was then used to inoculate one litre of LB-carbenicillin medium and was subsequently incubated at 37°C and 180 rpm.

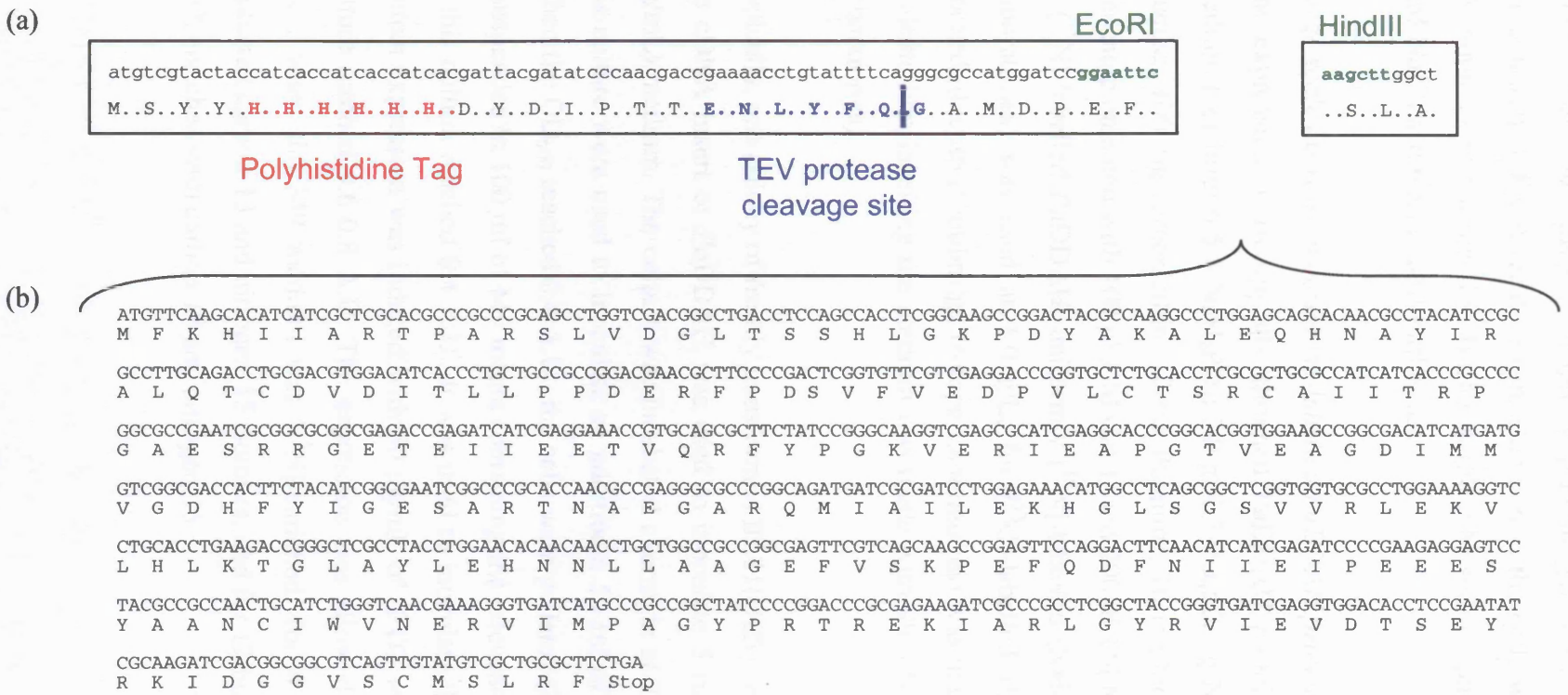

Figure 2.1. (a) DNA and amino acid sequence of the *Eco*RI and *Hind*III cloning sites of pPROEX Hta (green). The polyhistidine tag (red) and the TEV cleavage site (blue) are indicated. (b) DNA and amino acid sequence of the *Pa*DDAH cDNA. The PaDDAH insert is situated between *Eco*RI and *Hind*III restriction endonuclease cleavage sites (the excised section of the multiple cloning site has been omitted).

When the optical density (OD) of the larger culture reached OD_{600} = 0.6-0.8 absorbance units (A.U.), the expression of *Pa*DDAH was induced for four hours at 30°C with 50 μg/mL isopropyl-β-D-thiogalactopyranoside (IPTG; Melford Laboratories, UK). After the expression period the cells were pelleted (6000 rpm for 15 minutes), resuspended in Buffer A (50 mM sodium phosphate buffer pH 8.0, 300 mM NaCl) and frozen until further use.

Large scale expression of isotope-labelled PaDDAH proteins

The expression of isotopically enriched *Pa*DDAH proteins was carried out in M9 medium (per litre: 6.5 g NaH_2PO_4; 3.0 g KH_2PO_4; 0.5 g NaCl; 1.0 g $(NH_4)_2SO_4$; 5 g glucose; 100 mg carbenicillin; 10 mg thiamine; 10 mg biotin; all filter sterilised and then supplemented with 100 μl 1 M $CaCl_2$ and 200 μl 1M $MgSO_4$). For the expression of [^{15}N]-labelled *Pa*DDAH, uniformly [^{15}N]-labelled $(NH_4)_2SO_4$ (Cambridge Isotope Laboratories) was used at 1.0 g/L; for [^{13}C]-labelled *Pa*DDAH, uniformly [^{13}C]-labelled glucose (Cambridge Isotope Laboratories) was used at 2 g/L; and in the case of deuterium labelling the medium was made with 99.5 % ^2H_2O (Cambridge Isotope Laboratories).

Typically, one colony of freshly transformed BL21(DE3) cells carrying the vector with the cDNA insert of *Pa*DDAH was used to inoculate 5 ml of LB-carbenicillin (100 μg/mL) medium. The culture was incubated overnight at 37°C and 250 rpm. 2 ml of this culture were used to inoculate an additional 50 ml of LB-carbenicillin medium. When the OD_{600} reached 0.4 A.U., the cells were pelleted (3000 rpm, 10 minutes) and resuspended in 100 ml of M9 media containing the relevant isotopes. When the OD_{600} of this culture reached 0.4 A.U., it was used to inoculate the final 1 L of M9 media. Protein expression was induced with 50 μg/mL of IPTG when the OD_{600} of the final culture reached 0.6-0.8 A.U. The expression was allowed to proceed at 30°C for 6 hours when the M9 medium was [^{15}N]-enriched, for 8 hours when the medium contained carbon-13 and nitrogen-15 isotopes, and for 12 hours when the cells grew in D_2O, enriched with carbon-13 and nitrogen-15.

Purification of PaDDAH proteins

The bacterial pellets obtained after protein expression were resuspended in 20 to 30 ml of Buffer A (50 mM NaH_2PO_4, 300 mM NaCl, pH 8.0) and the suspension was passed three times over a French press (American Instrument Company). The cell debris was separated by centrifugation at 20,000 rpm for 30 minutes. The supernatant containing the soluble protein was then incubated for one hour at 4°C with nickel-charged nitriloacetic acid His·Bind® resin (Ni-NTA (Novagen); approximately 2.5 ml resin/500 ml bacterial culture). The purification procedure included three consecutive washes of the resin with 3×10 ml of Buffer A and three additional washes of 10 ml each of Buffer A with 10 mM imidazole. Each wash was followed by a quick centrifugation step to separate the supernatant from the resin (2500 rpm, 2 minutes). All supernatants were kept. The elution of the protein was achieved by incubating the resin with 4×10 ml of Buffer B (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, pH 8.0). Each wash and elution step was monitored by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) (See Figure 2.2a for a representative SDS-PAGE gel of the fractions obtained from the affinity chromatography step). The elution fractions containing the *Pa*DDAH enzyme were combined for the removal of the affinity tag. Recombinant tobacco etch virus protease (rTEV; also containing an N-terminal polyhistidine tag) was added to *Pa*DDAH at a ratio of 1:20. The reaction mixture was then dialysed against 2 L of Buffer A for 48 hours at 4°C. The cleavage reaction was monitered by SDS-PAGE (Figure 2.2b) and the cleaved *Pa*DDAH was separated from both rTEV and uncleaved *Pa*DDAH by a second Ni-NTA affinity chromatography step. Fractions containing the cleaved protein were concentrated to a final volume of 6 to 12 ml with Vivaspin20® sample concentrators. *Pa*DDAH was further purified with preparatory-scale Superdex 75 (S75) or Superdex 200 (S200) size exclusion chromatography (SEC) using Buffer C (20 mM sodium phosphate buffer pH 7.0, 100 mM NaCl, 1mM EDTA). SEC columns were attached to a Pharmacia FPLC (Amersham Biosciences, UK) or a BIOcad Sprint perfusion system (Applied Biosystems) and washed at a flow rate of 0.5 ml/min. Protein elution was followed by absorbance at 280 nm (Figure 2.2c). *Pa*DDAH-containing fractions were confirmed by SDS-PAGE (Figure 2.2d). Typically protein expression in either 2 L of LB or M9 media yielded approximately 25-30 mg of purified cleaved *Pa*DDAH protein.

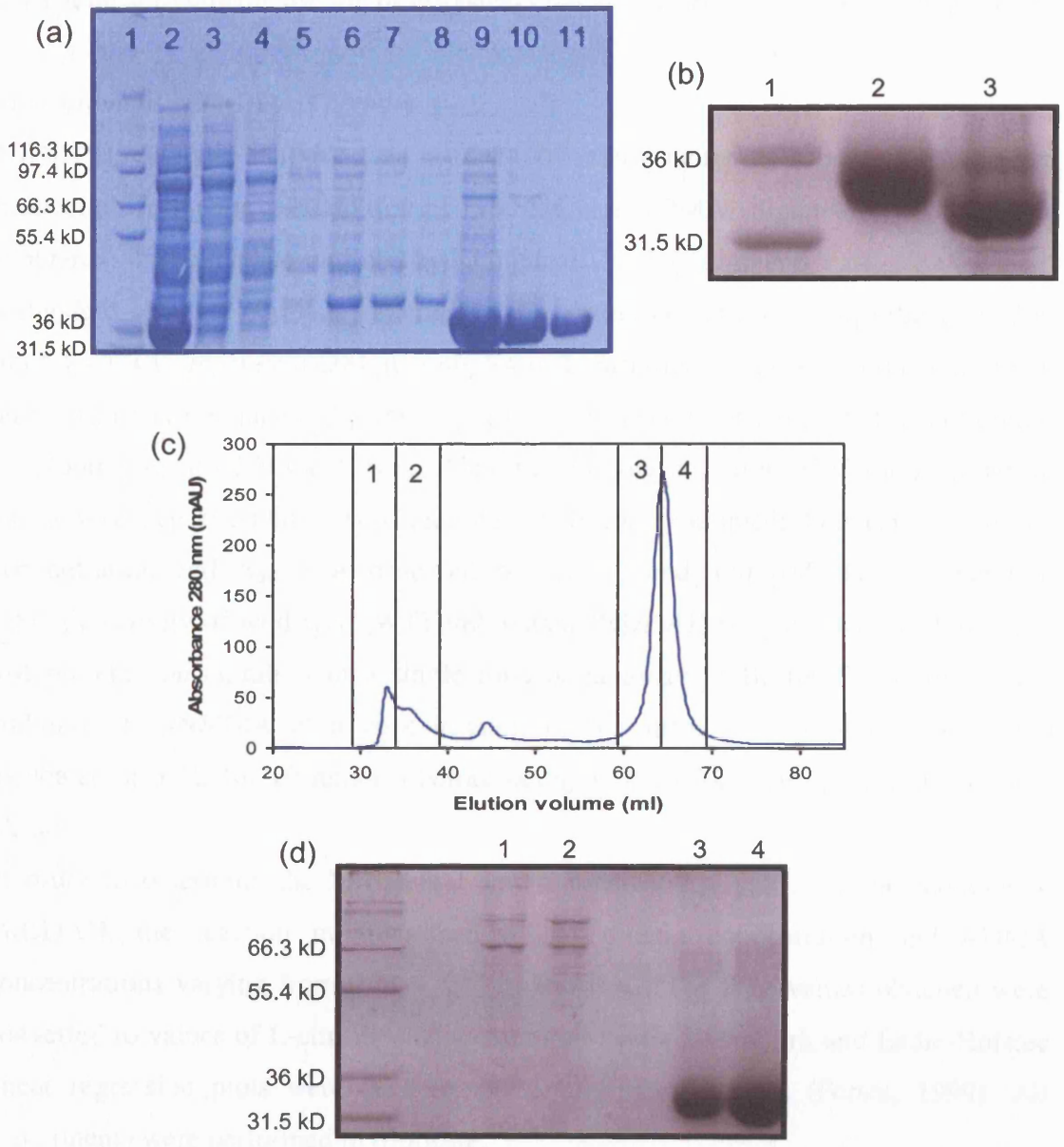

Figure 2.2. The purification strategy employed for R40E-R98H *Pa*DDAH. (a) SDS-PAGE analysis of affinity chromatography fractions of *Pa*DDAH using His-Bind® resin (Novagen). Lanes from left to right - 1: Mark12® protein molecular weight standards (Novex); 2: post lysis cell extract; 3-5: washes with Buffer A (See Chapter 2, Section 2.3); 6-8: washes with Buffer A plus 10 mM imidazole; 9-11: elutions with Buffer B (250 mM imidazole). Poly-His tag fused *Pa*DDAH elutes between 36 and 31.5 kD markers (indicated). (b) Removal of the polyhistidine tag by rTEV protease cleavage. From left to right – 1: as in (a); 2: concentrated elutions 9 to 11 from (a); 3: De-tagged *Pa*DDAH after 72 hour cleavage reaction with rTEV. De-tagged *Pa*DDAH runs to approximately 32 kD in SDS-PAGE analysis. (c) An elution profile of cleaved *Pa*DDAH following preparative scale gel filtration chromatography using a preparative scale S200 gel filtration column (See Chapter 2. Section 2.3). Vertical black lines indicate 5 ml fractions collected, 1 to 4. (d) SDS-PAGE analysis of fractions 1 to 4 from (c).

2.4 General protocols for the biophysical characterisation of *Pa*DDAH proteins

Measurement of PaDDAH activity

*Pa*DDAH catalytic activity was assayed by colorimetric detection of L-citrulline production following the addition of L-NMMA or ADMA (Sigma-Aldrich) using an adaptation of the method outlined by Knipp and Vašák (Knipp and Vašák, 2000). First, a standard curve to correlate L-citrulline concentration and optical absorbance at 540 nm (A_{540}) was generated: 200 µL samples of L-citrulline at concentrations between 0 and 2 mM were prepared (3 samples per L-citrulline concentration). 666 µL of colour developing agent (COLDER) was added, the mixture heated at 95°C for 15 minutes, and absorbance at 540nm (A_{540}) recorded. A linear relationship between L-citrulline concentration and A_{540} was observed between 0 and 400 µM. The comparative catalytic activity of wild-type (WT) and mutant *Pa*DDAH enzymes was analysed at 4 µM protein concentration in a single time point assay in Buffer C. Following the addition of L-NMMA at a concentration of 5.5 mM, the reaction mixture was incubated at 37°C for 10 minutes before being assayed for L-citrulline concentration (A_{540}).

In order to determine the Michaelis-Menten constant K_M value for the monomeric *Pa*DDAH, the reaction mixtures had 4 µM protein concentration and ADMA concentrations varying from 0 to 3 mM in Buffer C. The A_{540} values obtained were converted to values of L-citrulline concentration. Lineweaver-Burk and Eadie-Hofstee linear regression plots were used in the determination of K_M (Fersht, 1999). All experiments were performed in triplicate.

Analytical size exclusion chromatography (SEC)

100 µL samples of protein at 5 to 0.05 mg/mL were loaded onto a 10/30 HR Superdex-75 (S75) analytical gel filtration column (Amersham Biosciences) pre-equilibrated with Buffer C. Protein elution was followed by absorbance at 280 nm at a flow rate of 0.5 mL/min. A Gel-Filtration Low Molecular Weight Calibration Kit (Amersham Biosciences) was used for the calibration of the elution volumes.

Analytical ultracentrifugation

Sedimentation equilibrium analytical ultracentrifugation (SE-AUC) experiments were performed using an Optima XL-I Beckman analytical ultracentrifuge (Beckman

Scientific, Palo Alto, USA). Wild-type and mutant *Pa*DDAH samples, with the His_6-tag removed by proteolysis, were prepared at concentrations ranging from 0.2 to 4 mg/mL in Buffer C. Samples were sequentially centrifuged at rotor speeds of 15,000, 19,000 or 20,000, and 24,000 rpm at 20°C until equilibrium was reached for each individual speed. An additional run at 42,000 rpm was performed to obtain the optical baseline offset. Interference data were analysed using XL-A/XL-I data analysis software in ORIGIN 4.1. The apparent weight-averaged molecular mass ($M_{w,app}$) was obtained by fitting the data to Equation 2.1, which assumes a single ideal molecular species.

$$C_r = C_{r_0} \exp\left[M_{w,app} \frac{\omega^2(1-\bar{\upsilon}\rho)}{RT}\left(r^2 - r_0^2\right)\right] \qquad \text{Equation 2.1}$$

C_r is the total protein concentration at radial distance r, C_{r0} is the concentration at a reference point r_0, R is the gas constant, T is the absolute temperature and ω the angular rotor speed. The specific partial volume, $\bar{\upsilon}$, for WT *Pa*DDAH and all *Pa*DDAH mutants was estimated using the program SLUV (Perkins, 1986). The solvent density, ρ, was calculated from the buffer composition using the program SEDNTERP (www.rasmb.bbri.org).

The self-association properties of *Pa*DDAH proteins were investigated by simultaneously fitting to a single global model up to nine interference data sets obtained from a range of cell loading concentrations and rotor speeds (see Appendix I). Protein concentration (C_r) *vs.* radius (r) plots were analysed to determine the association constant K_{a2} (in units of fringes^{-1}, $K_{a2} = K_{fringe}$) of WT *Pa*DDAH and *Pa*DDAH mutants according to the analytical relationship predicted for a monomer-dimer equilibrium (Equation 2.2).

$$C_r = C_{r_0} \exp\left[M \frac{\omega^2(1-\bar{\upsilon}\rho)}{RT}\left(r^2 - r_0^2\right)\right] + \frac{2K_{a2}}{M}\left(C_M r_0\right)^2 \exp\left[2M \frac{\omega^2(1-\bar{\upsilon}\rho)}{RT}\left(r^2 - r_0^2\right)\right]$$

Equation 2.2

where C_M is the concentration of monomeric species at the cell radius r_0, and M is the monomer molecular mass, R is the gas constant, and T is the absolute temperature

(McRorie and Voelker, 1993). The association constant K_{a2} or K_{fringe} obtained from the best fit non-linear regression is converted to molar units using Equation 2.3:

$$K_{conc} = K_{fringe}\left(\frac{\frac{dn}{dc}\cdot l}{\lambda}\right)^{n-1}\frac{(M)^{n-1}}{n} \qquad \text{Equation 2.3}$$

where K_{conc} is the association constant in molar concentration, K_{fringe} is the signal association constant, *dn/dc* is the specific refractive increment (L/g), l is the centrepiece pathlength (cm), λ is the lightsource wavelength (cm), M is the monomer molecular weight and n the maximum stoichiometry of the association. Most proteins have a consistent value of (dn/dc × l/λ) of 3.31 fringes/g/L, when assuming a wavelength of 675 nm for the laser diode, a 1.2 cm pathlength and a specific refractive increment of 0.186 ml/g.

2.5 Nuclear magnetic resonance spectroscopy

Spectrometers

NMR spectra were recorded at the UCL NMR laboratory on Varian UnityPlus spectrometers at 500 and 600 MHz proton frequency equipped with triple resonance Z-axis pulse field gradient (PFG) probes with four and three radio frequency (RF) channels respectively. All pulse sequences employed on UCL spectrometers had been previously adapted by staff members and the experiments were set up by the author alone or under the guidance of Dr R. Harris.

All ^{15}N relaxation experiments were recorded at the MRC Biomedical NMR centre, Mill Hill (NIMR), London on Varian spectrometers unless otherwise stated, operating at proton frequencies of 500, 600 and 800 MHz. The 800 MHz spectrometer is equipped with four RF channels, single-axis pulsed field gradients and with a triple-resonance PFG probe. The experiments recorded at the NIMR were performed by Dr T. Frenkiel.

Heteronuclear single quantum coherence (HSQC)

[^{1}H, ^{15}N]-HSQC experiments were recorded with pulse field gradient sensitivity enhancement and with the incorporation of water flip back pulses for optimal water suppression (Zhang *et al.*, 1994)

Transverse relaxation optimised spectroscopy (TROSY)

[^{1}H, ^{15}N]-TROSY pulse sequences were downloaded with permission from Prof. L.E. Kay (http://pound.med.utoronto.ca/pulse_reg.html) at the University of Toronto, Canada and modified in house by Dr R. Harris to incorporate the WATERGATE pulse sequence element for water suppression (Piotto *et al.*, 1992).

NMR triple resonance experiments for sequence specific resonance assignment

All 2D and 3D NMR spectra acquired for sequence specific resonance assignment were recorded with a 1 mM [^{2}H,^{13}C,^{15}N]-R40E-R98H *Pa*DDAH sample at 25°C and at a proton frequency of 500 MHz. The following experiments were acquired with established pulse sequences: HNCA, HN(CO)CA, HN(CA)CB, HN(COCA)CB, HNCO (Yamazaki *et al.*, 1994a). The parameter details of each experiment are listed in Table 2.2. Sequence specific resonance assignments were found by correlating intra- and inter-residue resonances of CO, C_α and C_β shifts in order to create connected segments of sequence, which were then assigned using the probabilistic method of Grzesiek and Bax (Grzesiek and Bax, 1993) based on the characteristic C_α and C_β shift distributions of each residue type.

Measuring ^{15}N longitudinal R_1, transverse R_2 relaxation rates and {^{1}H}-^{15}N nuclear Overhauser effects (NOE)

^{15}N R_1, R_2 and {^{1}H}-^{15}N nuclear Overhauser effect (NOE) experiments were recorded using experiments based on published pulse sequences (Kay *et al.*, 1989) and modified as shown in Chapter 5, Section 5.2. The samples of the apo-form of ^{15}N-labelled R40E-R98H *Pa*DDAH were prepared at 1 mM concentration in Buffer C. The relaxation delays and the pre-saturation periods are detailed in Table 2.3. Spectra were acquired at 26°C by using 1440 x 200, 2000 x 240 and 2048 x 360 complex points in the t_2 x t_1 dimensions at ^{1}H proton frequencies of 500, 600 and 800 MHz respectively, with 16-48 scans per t_1 increment, a recycle delay of 1.4 s in R_1 and R_2 experiments, and a recycle delay of 1.2 s added prior to the ^{1}H saturation period in the {^{1}H}-^{15}N

Table 2.2. Experimental details for 2D and 3D NMR experiments acquired on [^{2}H,^{13}C,^{15}N]-R40E-R98H *Pa*DDAH at 25°C[a]

	^{1}H		^{13}C		^{15}N		
Experiment	Points	Sweep Width	Points	Sweep Width	Points	Sweep Width	Transients
HSQC	1024 (116ms)	4400			256 (150ms)	1700	8
HNCA	1024 (122ms)	4200	64 (18ms)	3600	32 (19ms)	1700	8
HN(CO)CA	1024 (122ms)	4200	64 (18ms)	3600	32 (19ms)	1700	8
HN(CA)CB	1024 (122ms)	4200	64 (9ms)	7100	32 (19ms)	1700	16
HN(COCA)CB	1024 (122ms)	4200	64 (8.5ms)	7500	32 (19ms)	1700	12
HNCO	1024 (122ms)	4200	48 (27ms)	1800	32 (19ms)	1700	8

[a]All spectra acquired on a Varian UnityPlus spectrometer operating at ^{1}H frequency of 500 MHz

Table 2.3. Delay periods for ^{15}N relaxation experiments of R40E-R98H *Pa*DDAH recorded at 26°C

Experiments	^{1}H frequency	Delay times for relaxation (s)[a]							
^{15}N T_1	500 MHz	0.0086	0.1210	0.2495	0.3940 (×2)[a]	0.6990	1.1000	1.6900	2.5000
	600 MHz	0.0086	0.1210	0.2495	0.3940 (×2)	0.6990	1.1000	1.6900	2.5000
	800 MHz	0.0086	0.1210	0.2495	0.4983 (×2)	0.9961(×2)	2.4974	2.9951	
^{15}N T_2	500 MHz	0.0087	0.0173	0.0347 (×2)	0.0520	0.0780	0.0953	0.1300	
	600 MHz	0.0086	0.0172	0.0344 (×2)	0.0516	0.0775	0.0947	0.1291	
	800 MHz	0.0085	0.0171 (×2)	0.0341	0.0682 (×2)	0.0767	0.0938	0.1194	
$\{^{1}H\}$-^{15}N NOE	500 MHz	0.000	3.0						
	600 MHz	0.000	3.0						
	800 MHz	0.000	3.0						

[a]In ^{15}N T_1 and ^{15}N T_2 experiments, the time periods stated refer to the length of time permitted for relaxation in the pulse sequence. In $\{^{1}H\}$-^{15}N NOE experiments the delay period refers to the length of the proton saturation period. [b]Indicates repeated times points.

NOE experiments. The spectral widths in the direct and the indirect dimensions were 7200 x 1670, 10000 x 2000 and 12000 x 3242 Hz at ^{1}H proton frequencies of 500, 600 and 800 MHz respectively.

An additional set of ^{15}N relaxation data was recorded at UCL for the 2-chloroacetamidine-bound *Pa*DDAH (1 mM protein concentration in Buffer C). The ^{15}N R_1, R_2 and {^{1}H}-^{15}N NOE relaxation parameters were measured at a ^{1}H proton frequency of 600 MHz at 26°C. The relaxation delays in seconds for the ^{15}N R_1 experiment were 0.010 (x2), 0.351, 0.702, 1.004, 1.405, 1.807 (x2) and for the R_2 experiments they were 0.015, 0.031 (x2), 0.062, 0.078, 0.935, 0.124. The ^{1}H saturation period in the {^{1}H}-^{15}N NOE experiment was 3.0 s. The spectral widths were 8000 x 1680 Hz in the direct and indirect dimensions. The spectra were acquired with 2048 x 256 complex points in the t_2 x t_1 dimensions with 26 to 46 scans per t_1 increment and a 2 s recycle delay during the R_1 and R_2 experiments. A recycle delay of 0.5 s was added to the saturation period during the {^{1}H}-^{15}N NOE experiment. The ^{15}N relaxation pulse sequences used in this set of experiments incorporated pulse field gradient sensitivity enhancement features (Kay *et al.*, 1992a) in contrast to the ones used at the NIMR (see Chapter 5, Section 5.2 for details).

Measuring ^{15}N R_2 Dispersion

^{15}N R_2 relaxation rates were only measured at ^{1}H proton frequency of 800 MHz with a relaxation-compensated–Purcell–Meiboom–Gill (CPMG) pulse sequence as described by others (Loria *et al.*, 1999; Tollinger *et al.* 2001). The CPMG period was implemented in a constant time manner, with T = 40 ms. Relaxation dispersion profiles were generated by measuring R_2 relaxation rates as a function of τ_{cp}, the time between successive 180° pulses in the CPMG sequence. Spectra were acquired as two-dimensional data sets at τ_{cp} intervals of 10, 5, 3.33, 2.5, 2, 1.66, 1.43, 1.25, 1 ms (data points 5 and 1 ms were duplicated). A reference spectrum with the CPMG blocks omitted was also obtained in duplicate. Spectra were recorded by using 2400 x 256 complex points in the t_2 x t_1 dimensions at 800 MHz with 64 scans per t_1 increment and 1.5 s recycle delay. The spectral widths were 12000 x 3242 Hz in the direct and indirect dimensions respectively.

NMR titration experiments

For NMR titration experiments various ^{15}N-labelled R40E-R98H *Pa*DDAH samples were prepared at concentrations ranging from 0.4 to 0.7 mM (Table 2.4). The ligands used in each titration experiment were provided by collaborators at the Department of Clinical Pharmacology at UCL, either in lyophilised form or dissolved in H_2O or DMSO. The concentrations of ligand reached at the end of the titration experiment are also shown in Table 2.4.

Table 2.4. Details of the NMR titration experiments performed with various ligands

Ligand	Total protein conc.[a] (mM)	Max. ligand conc.[b] (mM)	Titrations points[c]	% Protein dilution[d]
ADMA	0.55	3.58	12	18
L-citrulline	0.58	5.93	14	10
SR 140	0.44	4.12	16	17
SR 227	0.70	10.4	20	20
SR 257	0.40	1.20	3	10
HB 65	0.56	1.68	3	8
HB 102	0.50	1.50	3	8

[a]Refers to the R40E-R98H *Pa*DDAH concentration at the beginning of the titration. [b]Refers to the maximal concentration of ligand reached at the last titration point. [c]Refers to the number of titration points acquired. [d]Refers to the % of protein dilution at the end of the titration experiment.

A comprehensive list of the structures of each titrated ligand can be found in Chapter 6, Section 6.3. ^{1}H-^{15}N HSQC experiments were collected for each titration point using either a 500 MHz or 600 MHz spectrometer. The spectral widths were 4200 x 1700 Hz and 5400 x 2100 Hz at 500 and 600 MHz respectively, with a total of 1024 x 192 complex points in the direct and indirect dimensions and 4-8 scans per t_1 increment. Prior to each acquisition the pH of the ligand and protein samples were painstakingly monitored and were adjusted to 7.0. The H^N and ^{15}N chemical shifts of selected cross peaks in each spectrum acquired were output for chemical shift mapping and the estimation of a dissociation constant K_d where possible (see Chapter 6, Section 6.2).

The binding curves were fit to a nonlinear regression analysis using the following Equation 2.4 to monitor the chemical shifts in the labelled protein resonance:

$$\delta_{obs} - \delta_{free} = (\delta_{sat} - \delta_{free}) \frac{(K_d + L_0 + P_0) - \sqrt{(K_d + L_0 + P_0)^2 - 4P_0L_0}}{2P_0} \quad \text{Equation 2.4}$$

where P_0 , L_0, are the total protein and ligand concentrations, and δ_{obs} and δ_{free} are the observed change in chemical shift on ligand titration and the chemical shift of the free state, respectively. δ_{sat} and K_d are to be determined by the regression analysis, where K_d is the dissociation constant and δ_{sat} is the chemical shift of the complex under saturating conditions. In these studies however, L_0 was substituted by $(L_0/P_e)*P_0$, where P_e is the protein concentration at each titration point.

2.6 NMR data analysis

Processing NMR data

All raw NMR data were processed using the nmrPipe program of Delaglio and co-workers (Delaglio *et al.*, 1995). Standard manipulations of data sets included zero-filling to the nearest 2^n points, application of window functions, base line corrections, linear prediction of indirect dimensions where necessary and zero and first order phase corrections were applied using nmrPipe. Spectra were initially visualised in nmrDraw (Delaglio, *et al.* 1995).

The processed spectra were exported into AZARA (http://www.bio.cam.ac.uk/azara/) format using the PIPE2AZARA command. Multiple spectra were visualised together and cross referenced in Plot2 (http://www.bio.cam.ac.uk/azara/). Sequential resonance assignment was performed using ANSIG (Kraulis, 1989) with contour and cross peak files generated by AZARA.

Derivation of ^{15}N R_1, ^{15}N R_2, ^{15}N $\{^1H\}$-heteronuclear-NOE parameters

^{15}N R_1 and R_2 experiments yielded a pseudo 3D-spectrum with the third dimension corresponding to the 2D NH correlation plane with incremented relaxation delays (See Table 2.3). For each set of relaxation data recorded, a reference 2D [^1H,^{15}N] spectrum was acquired. Only well resolved NH cross peaks in the reference spectrum were selected for further analysis. The signal intensity of each selected cross peak from the 2D planes recorded was extracted using ANSIG scripts written in-house (Pfuhl *et al.*, 1999). The peak intensity values for a single cross peak were output as a function of relaxation delay. The decay of intensity was fit to Equation 5.7 (see Chapter 5, Section 5.2 for more details) with the *Mathematica*® software (Wolfram Research Inc.) using a two parameter least square fit (Levenberg-Marquardt method as described in Kay *et*

al., 1992b; Peng and Wagner, 1994). The error on the fit was estimated using a Monte Carlo simulation based on the cross peak signal intensities acquired for a pair of duplicated time points. The relaxation rates and standard deviations stated in the text correspond to the mean relaxation rate and standard deviation of the distribution of two hundred such Monte Carlo iterations.

$\{^1H\}$-^{15}N heteronuclear NOE experiments were acquired as a pair of 2D NH correlation spectra in the presence and absence of proton saturation at three different magnetic fields (Table 2.3). The signal intensity values of selected cross peaks were extracted as described for the R_1 and R_2 relaxation experiments (See Chapter 5, Section 5.2 for more details). The steady-state NOE values were determined from the ratios of the intensities of the peaks with and without proton saturation. The experimental error σ_{NOE} was calculated based on the spectral noise value (σ_{Isat} and σ_{Iunsat}) in each 2D NH plane using the following relationship (Equation 2.5):

$$\sigma_{NOE} / NOE = ((\sigma_{I_{sat}} / I_{sat})^2 + (\sigma_{I_{unsat}} / I_{unsat})^2)^{1/2} \qquad \text{Equation 2.5}$$

where I_{sat} and I_{unsat} are the measured intensities of a cross peak in the presence and absence of proton saturation, respectively.

Chapter 3

The engineering of a monomeric *Pa*DDAH

Abstract

Proteins capable of self-association benefit from an increased stability and the ability to perform additional functional tasks through intersubunit communication. Despite the large number of protein oligomers present in nature, a significant proportion of the oligomeric structures available are the result of non-specific crystallographic contacts caused by the higher protein concentrations required for structural characterisation. The efforts to study the 58 kD homodimeric enzyme *Pa*DDAH by NMR were hampered by the molecular size of the molecule which led to substantial line broadening. With the crystal structure of *Pa*DDAH available, it became increasingly clear that the small interface surface area of the dimer and the fact that conditions could be found where *Pa*DDAH exhibited a monomer-dimer equilibrium in solution were indicative of the non-obligate homodimeric nature of the enzyme. This chapter describes the exploration of the interface region of *Pa*DDAH through the careful analysis of the residues responsible for self-association. The analysis led to the design of a series of interface mutants with the ultimate goal to generate a stable active monomeric enzyme. These mutants were extensively characterised by size-exclusion chromatography, analytical ultracentrifugation and NMR spectroscopy. Dissociation constants were derived for a selected set of *Pa*DDAH mutants in order to quantify effectively the disruption of the association. One stable and active double mutant R40E-R98H *Pa*DDAH behaved as an exclusive monomer in solution and yielded superlative 2D [^{1}H,^{15}N]-HSQC NMR spectra when compared to wild-type *Pa*DDAH. Our findings confirm the initial observations that the dimerisation of *Pa*DDAH is not a prerequisite for biological activity and that the pursuit of a monomerisation strategy was well founded in order to suppress the dynamics of the monomer-dimer equilibrium of the protein and decrease the overall size of the protein, thereby providing a form of DDAH that is optimal for further analysis by NMR.

3.1 Introduction

Functional role and evolution of protein oligomerisation

Proteins commonly act as oligomers in cells. A survey of the SWISS-PROT database for *Escherichia coli* annotated proteins shows that dimers and tetramers constitute up to 60% of the 617 entries with explicit "subunit" annotations, whilst monomeric proteins represent only 19% of the total (Goodsell and Olson, 2000). The structural and functional advantages conferred to proteins by oligomerisation include the opportunity for cooperative interaction and allosteric regulation between subunits (Marianayagam *et al.*, 2004). Oligomerisation has also been implicated in the regulation of enzyme function (Traut, 1994). Several enzymes are reported to self-assemble in order to promote catalysis, a classic example of which is *Escherichia coli* aspartate transcarbamoylase (Wang *et al.*, 2005). In this respect, self-association will only occur if a threshold protein concentration is reached within cells, so that subunit assembly is favoured. Moreover, the stable association of protein subunits is required in order to produce higher order structural scaffolds such as viral capsids (Johnson and Speir, 1997). The expression and assembly of identical subunits has even been suggested to act as a mechanism enabling the reduction of the viral genome size. Another advantage conferred by protein self-association is an increase in the overall stability of the oligomer by the cumulative enthalpic contributions of hydrogen bonds, van der Waals forces and hydrophobic interactions present across subunits (Jaenicke, 2000). An illustration of increased stability through oligomerisation has been suggested in thermophilic enzymes that associate as dimers, whilst their mesophilic counterparts are monomeric and more prone to denaturation at higher temperatures (Thoma *et al.*, 2000).

To date, all models explaining the evolution of oligomerisation share the precept that a monomeric form of the protein pre-dates multimer formation. The "mutation model" put forward by D'Alessio and colleagues (D'Alessio, 1999) argues that a primary mutational event on the surface of a monomeric protein is responsible for rendering the mutant less stable and susceptible to adhere to another molecule. In thermodynamic terms, this mutational event translates into a decrease in the free energy barrier that separates free monomers from associated oligomers. Additional mutations can also occur and serve to strengthen the subunit association. Although it is difficult to assess

what the primary mutational event might have been from the visual inspection of the 3D structures of present day oligomers, several studies of directed *in vitro* evolution have generated engineered dimers from monomeric *Staphylococcal* nuclease by the simple deletion of residues in a surface loop (Green *et al.*, 1995).

A second evolutionary model for dimerisation proposed by Xu *et al.* (1998) states that gene mutation, deletion, and fusion of two chains are pathways to obtain foldable and functional proteins without the need for an intermediate stable monomeric precursor. This hypothesis is derived from observations of oligomers that show dimerisation kinetics that resemble the folding kinetics of monomers.

The third pathway is the "swap model" in which the emergence of oligomerisation occurs in a single step (Bennett *et al.*, 1994; 1995). A dimeric protein evolves from an interchange of C- or N-terminal secondary structural elements of two monomeric subunits. This domain swap creates an interdomain interface that does not require random mutations to occur in the exposed surface area of a protein. However, D'Alessio argues that such a model should also be subject to mutational events which might be responsible for disrupting the favourable interactions that held the secondary structural elements within their original subunits. In this respect, such a model should not be classified as an evolutionary pathway for oligomer formation, but rather as a mechanistic approach to explain the monomer to oligomer transition.

General trends of oligomer interface surfaces

The common assumption held from the studies of known X-ray structures of oligomers is that hydrophobic contacts in the interface are the driving force for oligomerisation in the same manner as the hydrophobic effect contributes to protein folding (Chothia and Janin, 1975; Honig and Nicholls, 1995). Consistent with this view are the findings by Nussinov and co-workers that show that the interface area is enriched by a moderate percentage (10-15%) in hydrophobic residues in comparison with their presence in the total surface area of oligomers (Tsai *et al.*, 1997). However, the novel view that emerged from their sampling of 362 interfaces is that the hydrophobic effect is not as crucial in protein-protein association as it is in protein folding, as polar and charged amino acids are more frequent in protein interfaces than they are in protein cores, although they remain predominant on monomer surfaces. This observation can be viewed as reflecting a compromise between achieving high stability of an interface with the presence of hydrophobic residues and at the same time maintaining the

stability of the independent folded subunit with polar and charged residues (Tsai *et al.*, 1997).

Larsen and colleagues corroborated these views by visually inspecting the structures of several protein interfaces, showing that in about 61% of cases, the interface shared small patches of hydrophobicity mixed with hydrophilic residues, hydrogen bonds and water molecules (Larsen *et al.*, 1998). Of particular relevance to this discussion are the statistical thermodynamic studies of Xu *et al.* that showed that in reality hydrophilic residue pairs contribute significantly to the binding free energy of monomers (Xu *et al.*, 1997), which is in contrast with what is observed in protein folding (Hendsch and Tidor, 1994). In general there will always be a destabilising effect to any assembly process due the loss of favourable solvation interactions between charged and polar residues and the high dielectric aqueous phase. The favourable electrostatic contribution to binding might include the entropic gain generated from the transfer of a hydrophilic ion-pair present on the surface exposed region of the protein to the buried interface domain of the associated state, overcoming the desolvation cost (Xu *et al.*, 1997). Desolvation effects are also partly overcome through the formation of networks of favourable Coulombic interactions between interfacial charged and polar side chains across or within subunits. In their studies, Honig and collaborators nicely illustrate that the greater the number of residues involved in inter-molecular hydrogen bonds and salt bridges with respect to the total number of charged and polar residues on the interface, the more favourable the total electrostatic contribution to binding will be (Sheinerman and Honig, 2002). These interfacial Coulombic interactions are strong enough to mimic the interactions that hydrophilic residues usually have with their aqueous solvation shell prior to protein oligomerisation.

The intriguing view that arises from the presence or absence of a mixture of hydrophobic and hydrophilic interactions in interfaces is that it does not correlate with any specific functional attribute. The proteins that have been sampled so far include enzymes, DNA-binding proteins, toxins and receptors, all of which can present a hydrophobic core upon oligomerisation or display a mixed interface with hydrophobic and hydrophilic regions (Larsen *et al.*, 1998).

Discriminating between functional oligomerisation and non-specific crystal contacts

The studies presented above, which refer to the description of subunit interfaces in terms of hydrophobicity and preference for certain amino acid residues, do not systematically contemplate the possibility that not all dimers or oligomers are required for protein activity. Much attention has been given lately to the development of protocols capable of discriminating between functional interfaces and non-specific crystal contacts in the unit cell. Initial qualitative findings have shown that the interfacial solvent accessible surface areas (ΔASA) are larger in biological contacts than in non-biological ones and usually constitute the largest lattice contact in the crystal (Jones and Thornton, 1996). Furthermore, since non-obligate oligomers can exist as independent molecular entities their interfacial composition tends to be analogous to the overall protein surface, hence more polar than in biologically relevant interfaces. Additionally, the interfacial contact surfaces of non-obligate oligomers are less complementary, as estimated by the Gap Index, a measure of the volume enclosed between any two interfaces (Jones and Thornton, 1996). Valdar and Thornton further extended these studies to combine the interface surface area and the conservation of amino acid residues in interfaces among homologous protein families to provide a robust prediction of biologically relevant dimers (Valdar and Thornton, 2001). They found that interfaces arising from artefacts of crystal packing rarely shared any amino acid conservation between homologues. Recent work by Bahadur *et al.* analysed 188 non-specific homodimers with large interface surface area (>800 $Å^2$), a value similar to that found for biologically specific interfaces. They showed that regardless of the size of the interface, on average physiologically relevant interfaces are more compact, better packed and contain more hydrogen bonds per polar atom than non-specific interfaces (Bahadur *et al.*, 2004).

Taken together, these findings provide guidelines for the effective discrimination between specific and non-specific interface contacts. The relevance of such studies often lies with examples of oligomers for which an experimental monomerisation strategy is sought. The availability of a monomeric protein entity can in most cases simplify the biophysical characterisation of a particular system and *a priori* knowledge of the type of interface that is targeted in such work certainly directs the mutagenesis strategy to efficiently disrupt the oligomer formation.

3.2 Analysis of *Pa*DDAH oligomeric state and interface surface area

PaDDAH hydrodynamic behaviour

The enzyme *Pa*DDAH is a 58 kD homodimer at concentrations necessary for structural characterisation (Murray-Rust et al., 2001). To successfully apply NMR methodology to the study of this enzyme, the challenge presented by the size of the homodimer had to be surmounted. Details of the problems encountered will be discussed in Chapter 4. Ultimately we had to redirect our attention to a site-directed mutagenesis strategy focused on the residues located at the dimer interface. Point mutations of these residues might weaken the association of the dimer without disrupting the active site region and result in the production of an active monomeric species more amenable to NMR investigations.

In the course of the pilot work on *Pa*DDAH, the first prompt for pursuing the monomerisation strategy came from the analysis of the hydrodynamic behaviour of *Pa*DDAH. On close inspection of the available data, the enzyme presented a monomer-dimer equilibrium in solution dependent upon protein concentration. Analytical size exclusion chromatography (see Chapter 2, Section 2.4) was the method of choice for characterising the self-associative properties of wild-type *Pa*DDAH (WT *Pa*DDAH). The enzyme elutes as a single asymmetric peak and shows subtle changes of elution volumes from 9.9 ml to 10.2 ml as the loading concentration decreased from 5 mg/ml to 0.05 mg/ml (Figure 3.1a,b). Conversion of the peak elution volumes into apparent molecular weights (MW_{app}) semi-quantitatively revealed a shift in the equilibrium towards monomeric species, from an initial *Pa*DDAH behaviour of a 50 kD species at high concentration, towards a $MW_{app} \approx 46$ kD at the lowest concentration (Figure 3.1c). For a protein with a predicted monomeric mass of 29 kD these data indicate a predominance of the homodimer in this concentration range.

A description of the protomer interface of PaDDAH

Building on these findings, the notion that a mutagenesis strategy could be pursued in order to obtain a monomeric variant of *Pa*DDAH was reinforced through the careful analysis of the interface residues involved in maintaining the associated homodimer. Inspection of X-ray structure predicts that the interface between the two protomers is composed of three secondary structural elements that form the contact points across each of the two polypeptide chains (defined here as protomers A and B; Figure 3.2a).

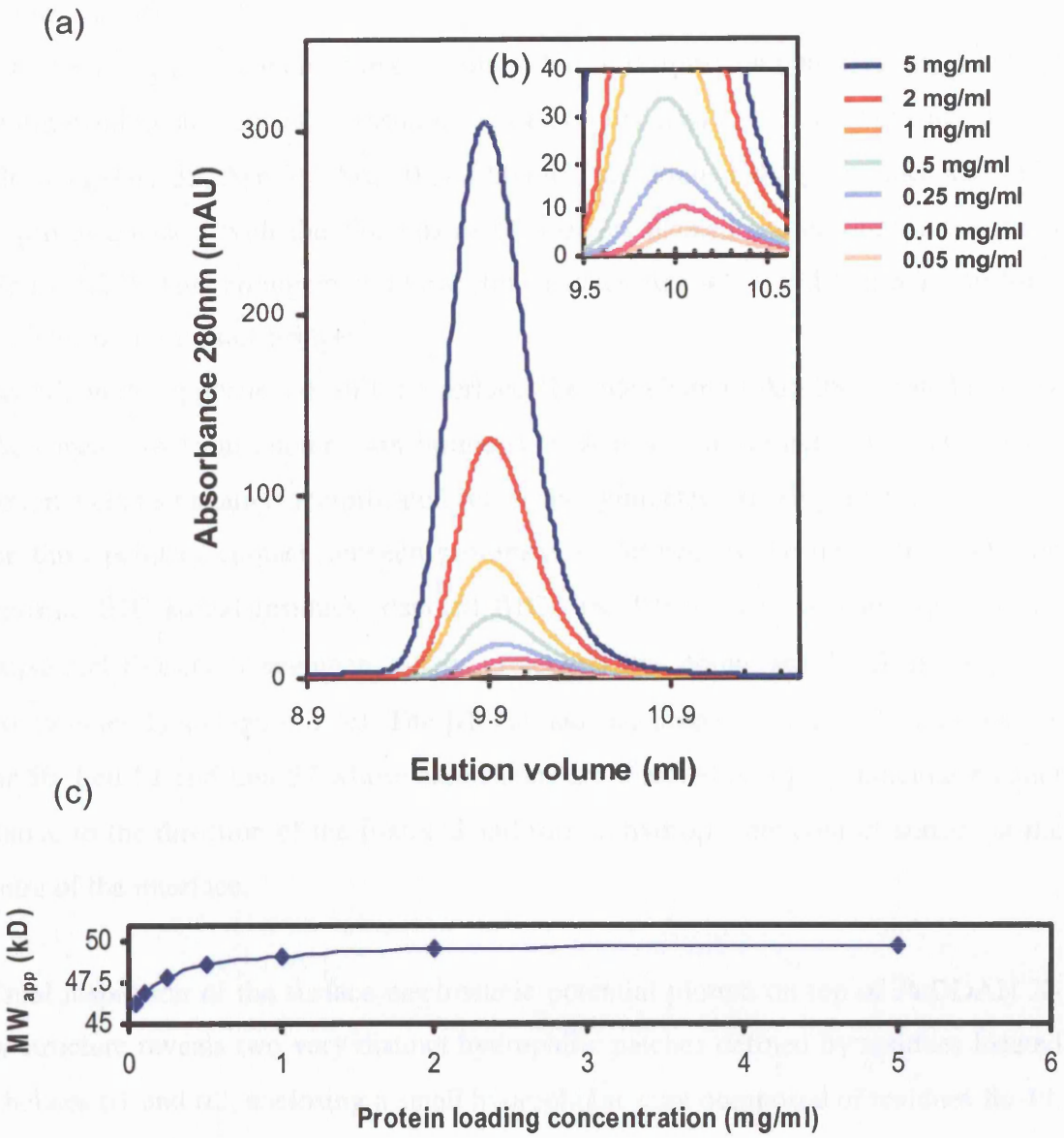

Figure 3.1. Analytical size exclusion chromatography of WT *Pa*DDAH. (a) Elution profile of WT *Pa*DDAH samples at different protein loading concentrations shown in the colour coded legend. (b) Blow-up of (a). (c) Apparent molecular weight (MW_{app}) of WT *Pa*DDAH derived from calibration of the SEC column with known MW standards as a function of protein loading concentrations.

The association is mediated by the side chain interactions of residues located in helices α1 and α2 of modules 1 and 2 and residues located in the β1C strand of module 1 (Figure 3.2b-d).

First, the side chain-side chain interactions between residues on opposing $\alpha 1_{A,B}$ helices are arranged in an antiparallel fashion, centred on residues Asn $36_{A,B}$. Four residues in helix-$\alpha 1_A$ (Glu 33, Asn 36, Arg 40 and Gln 43) are aligned along one face and form reciprocal contacts with the side chains of residues located on the opposing subunit (Figure 3.2d). This arrangement allows for residues Arg $40_{A,B}$ and Glu $33_{B,A}$ to form two inter-protomer salt-bridges.

Second, on the opposite side of the interface, the side chain of Arg 98_B located in helix $\alpha 2_B$ appears to form another salt-bridge with Asp 48_A of strand $\beta 1C_A$. This inter-protomer contact again is reciprocated across the symmetry axis (Figure 3.2b).

The third point of contact between protomers is defined by the interaction between opposing β1C strand residues (denoted $\beta 1C_A$ and $\beta 1C_B$). The two strands form an antiparallel β-sheet arrangement across protomers that comprises the β-sheets of the two modules $1_{A,B}$ (Figure 3.2c). The β1C strands encompass residues Asp 48, Ile 49, Thr 50, Leu 51 and Leu 52 whose side chains are directed in a perpendicular manner relative to the direction of the β-strand and form a hydrophobic contact surface at the centre of the interface.

Visual inspection of the surface electrostatic potential plotted on top of *Pa*DDAH X-ray structure reveals two very distinct hydrophilic patches defined by residues located in helices α1 and α2, enclosing a small hydrophobic core composed of residues Ile 49, Leu 51 and Leu 52 of the β1C-strand (Figure 3.3a,b). To examine whether this interface arrangement was conserved in other enzymes of the same superfamily, the interfaces of two arginine amidinotransferases AGAT and StrB1 were analysed in the same manner (Figure 3.3c-f). These two enzymes form homodimers in solution and the secondary structural arrangement in the interfacial region is strikingly conserved (Humm *et al.*, 1997, Fritsche *et al.*, 1998). The secondary structural elements involved in the interaction are part of the β-propeller modules 1, 2 and 3 and are centred about the 3-stranded β-sheets of module 2 (Figure 3.3c-f). However, these β-sheets do not extend across subunits as observed in the *Pa*DDAH interface. Instead the two opposing

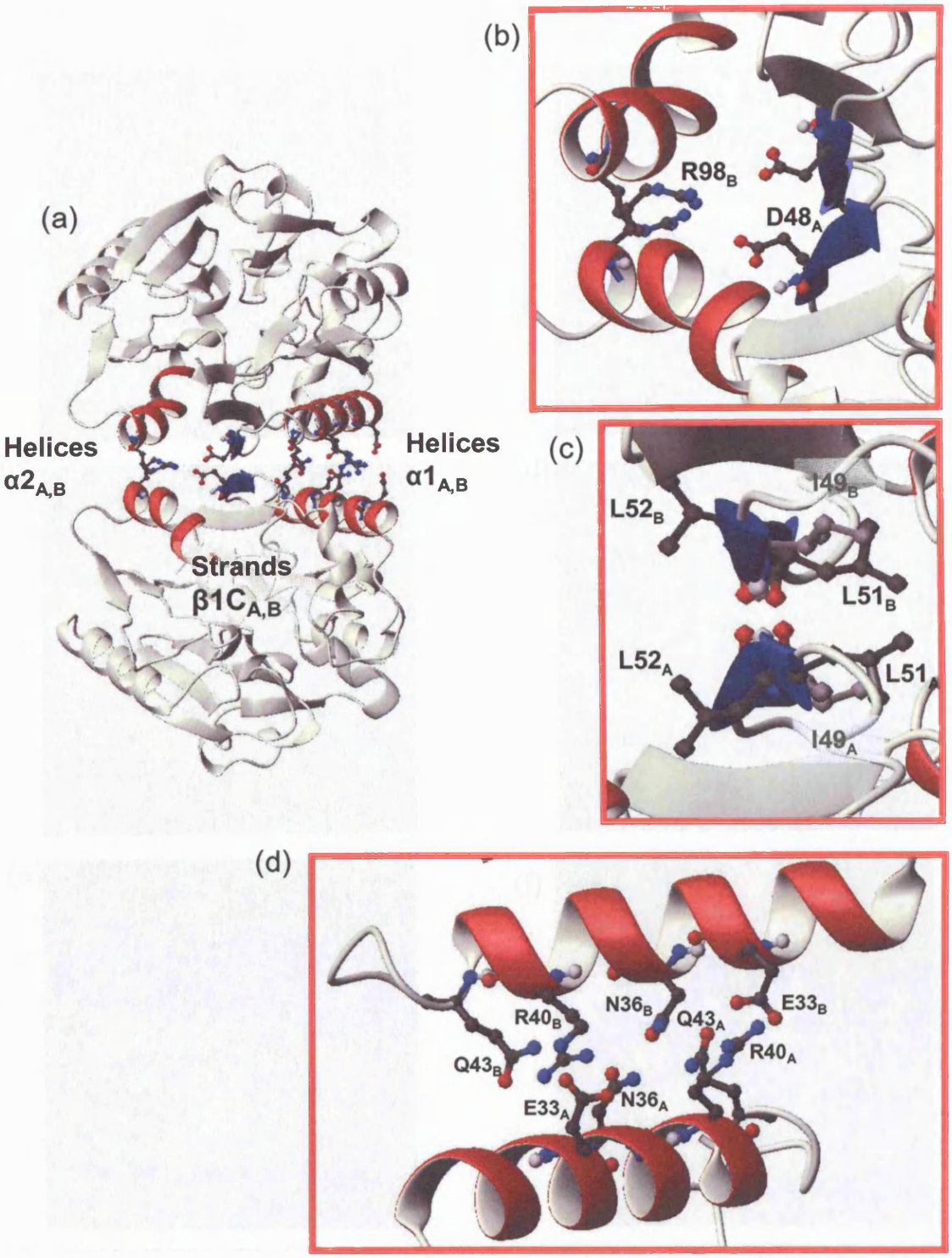

Figure 3.2. (a) The 3D structure of *Pa*DDAH homodimer interface, highlighting in red, helices $\alpha1_{A,B}$ and $\alpha2_{A,B}$ and in blue, strands $\beta1C_{A,B}$. (b) Close-up of the interaction between residues $R98_B$ in helix $\alpha2_B$ and $D48_A$ in strand $\beta1C_A$. The interaction is reciprocated across the symmetry-related protomers. (c) Detail of the hydrophobic interactions occurring between residues I49, L51, L52 side chains across strands $\beta1C_{A,B}$. (d) Interaction between side chains occurring across helices $\alpha1_{A,B}$. All atoms follow CPK colouring. Figures were created with MOLMOL (Koradi *et al.*, 1996).

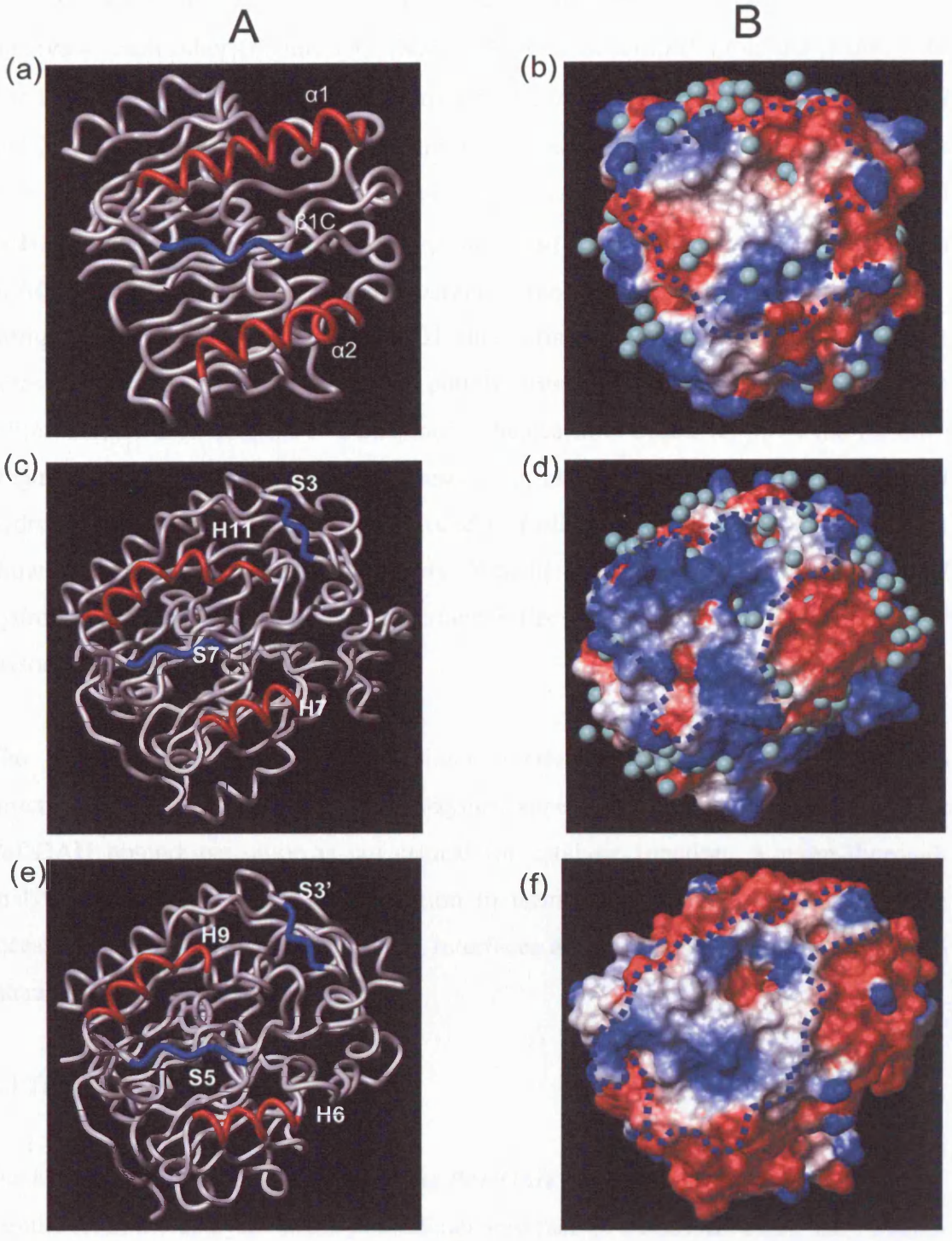

Figure 3.3. (A) Worm diagrams of monomeric (a) *Pa*DDAH (1h70), (c) AGAT (1jdw) and (e) StrB1 (1bwd), indicating the orientation of the molecule used in part (B). α-helices involved in the interface are red and β-strands are blue. The numbering of the helices (H) and strands (S) in (c) and (e) is taken from Fritsche *et al.* (1998). (B) 3D molecular surface of (b) *Pa*DDAH, (d) AGAT, (f) StrB1 coloured according to residue electrostatic potential. The interface surface is delimited by the blue dotted line. Residues in blue are positively charged, red negatively charged, white are hydrophobic and uncharged. Water molecules when present in the pdb files are depicted as pale green oxygen atoms. Figures were created with MOLMOL (Koradi et al., 1996).

sets of strands (strands $S7_{A,B}$ in AGAT and strands $S5_{A,B}$ in StrB1) form an angle relative to each other (Figure 3.4). Despite the lack of an inter-protomer β-sheet, the interactions of AGAT and StrB1 conferred by the α-helices and β-sheets of modules 2 and 3 are structurally homologous to those of modules 1 and 2 of *Pa*DDAH (Figure 3.4). An additional module is involved in the interfacial interaction of AGAT and StrB1 – module 1 – positioning two opposing 2-stranded β-sheet loops (strands $S3_{A,B}$ in AGAT and strands $S3'_{A,B}$) in an antiparallel arrangement (Figure 3.4).
Intriguingly, although AGAT and StrB1 share similar secondary structural elements across interfaces, the analysis of the putative residue electrostatic potential of the buried surface area of these two enzymes indicates little conservation of the residues (Figure 3.3B). AGAT appears to possess a clear mixture of hydrophilic and hydrophobic exposed side chains compared to *Pa*DDAH and StrB1 interfaces which show distinct patches of hydrophobicity. Whether the distributions of charges and hydrophobic sites on the surface interface reflect a difference in the homodimer dissociation constants is unknown.

The non-conservation of the homodimer interface structure in these otherwise structurally and functionally related enzymes appears consistent with the idea that *Pa*DDAH homodimerisation is not critical for catalytic function. A more thorough analysis is performed in the next section in terms of the actual values of surface accessible surface areas of the proteins' interfaces and the residues responsible for the interaction between protomers.

3.3 Targeting residues for mutation

The solvent accessible surface area of the PaDDAH interface

Results from the analysis of the homodimer interface of *Pa*DDAH using the Protein-Protein Interaction (PPI) server (Jones and Thornton, 1996) have shown that the interface surface area occluded from a solvent molecule with a radius of 1.4 Å upon dimer formation is approximately 860 $Å^2$ per subunit. This value and several other parameters believed to define the characteristics of an oligomer interface have been calculated for crystallographic homodimers of *Pa*DDAH, AGAT and StrB1 (Table 3.1). In their first review of protein-protein interactions which evaluated 72 homodimer

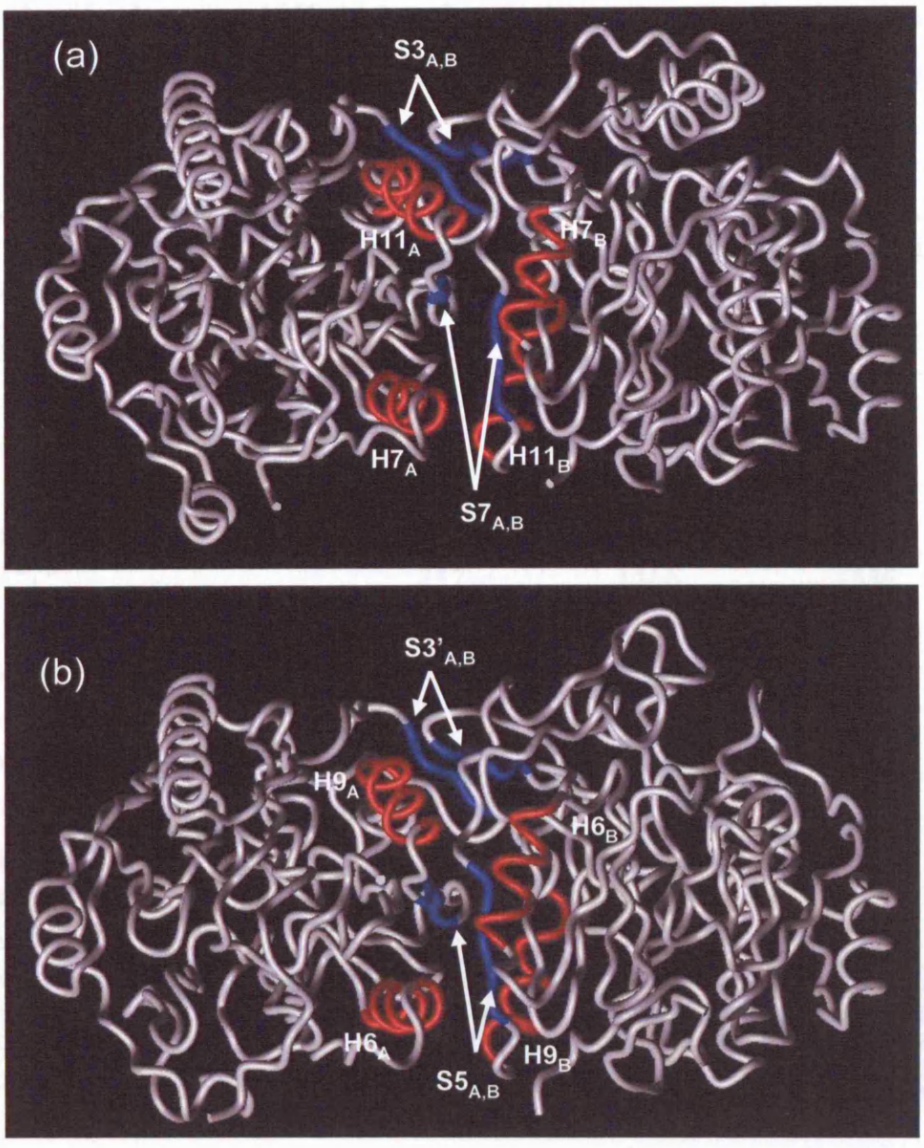

Figure 3.4. Worm diagrams of the 3D structures of the homodimeric (a) AGAT (1jdw) and (b) StrB1 (1bwd). The secondary structure elements involved in the interface are indicated: the α-helices are coloured in red and the β-strands are shown in blue. The numbering of the helices and strands in (a) and (b) is taken from Fritsche *et al.* (1998); S stands for β-strand and H for α-helix and the subcript letters A and B refer to each protomer. Figures were created with MOLMOL (Koradi *et al.*, 1996).

interfaces, Jones and Thornton established a weak linear correlation (coefficient ≈ 0.69) between the size of the interface surface area ($Å^2$) and the molecular mass of the monomer (Daltons). From their regression analysis the *Pa*DDAH interface area would have a predicted value of *ca.*1650 $Å^2$ based upon the monomer molecular mass of 29 kD and similarly, AGAT and StrB1 should have had a much larger interface areas than they actually do.

.

Table 3.1. Analysis of the homodimer interfaces from the crystal structures of WT *Pa*DDAH, AGAT and StrB1 using the Protein-Protein Interaction server[a]

Parameter	Protein		
	*Pa*DDAH	AGAT	StrB1
PDB Code	1h70	1jdw	1bwd
Number of Residues	254	423	348
Protomer Mass (kD)	29.0	48.5	39.0
ΔASA ($Å^2$) [b]	857	1466	1556
ΔASA as % of total ASA [c]	7.8	10.9	11.0
Residues with ΔASA > 0.0 $Å^2$	23	35	39
% Polar atoms in interface	49.6	35.8	31.7
% Non-polar atoms in interface	50.4	64.1	68.2
Hydrogen bonds	8	18	13
Salt bridges	2	2	0
Gap Volume ($Å^3$) [e]	5709.5	4782.2	5487.9
Gap Index (Å) [f]	3.33	1.63	1.76

[a]http://www.biochem.ucl.ac.uk/bsm/PP/server/; [b]ΔASA is the total ASA per protomer concealed on the formation of the homodimer; [c]ΔASA is the buried surface expressed as a percentage of the ASA of the monomer; [d]calculated with the program HBPLUS (McDonald and Thornton, 1994); [e]calculated with the program SURFNET (Laskowski, 1995); [f]GAP Index = Gap Volume/ 2ΔASA.

In their more recent review, Ponstingl and co-workers have shown that a power law better describes the correlation found between the monomer molecular mass and the interface area of the homodimer (Ponstingl *et al.* 2005). Their investigations based upon 76 protein homodimers revealed that on average 15% of the total accessible surface area of the subunit is buried upon dimer formation. For an additional set of 56 dimers termed "dummy dimers" to refer to self-associative proteins thought to exist only under crystallographic conditions, they found that this average falls below 10%. In the case of *Pa*DDAH, AGAT and StrB1, for which the percentage of total subunit ASA buried is 7.8%, 10.9% and 11% respectively, it appears as though these proteins might be non-obligatory dimers.

The same authors have also been able to determine that the hydrophobicity, defined as the fraction of carbon atoms present in the interface area, is an additional parameter capable of discriminating between obligate dimer and non-specific dimer interfaces. The fraction of carbon atoms contributing to the interface ASA in dummy interfaces is around 50% whereas in definite homodimers it increases to an average of 62%. However, a more cautious analysis shows a rather large variance for monomers with smaller crystal contacts, as several dimer interfaces in this class were found to have carbon atoms covering up to 85% of the interface area. This assessment could explain the high percentage of non-polar atoms found in the interfaces of AGAT and StrB1 (64% and 68% respectively, Table 3.1). For *Pa*DDAH the low percentage of non-polar atoms in the interface clearly places this protein into the non-obligatory homodimer class as predicted by this analysis.

Another key parameter in the definition of an oligomer interface is the topological complementarity between surfaces. Jones and Thornton proposed the calculation of a Gap Index, defined in Equation 3.1, where the gap volume is the empty volume between two protomers calculated with the program SURFNET (Laskowski, 1995) and the total ΔASA is the total surface area buried upon oligomerisation.

$$\text{Gap Index (Å)} = \frac{\text{Gap Volume (Å}^3\text{)}}{\text{Total } \Delta\text{ASA (Å}^2\text{)}} \qquad \text{Equation 3.1}$$

In the original work on this topic, an average Gap Index of 2.2 ± 0.9 Å was obtained from the analysis of 32 homodimers (Jones and Thornton, 1996). Despite the large Gap Index value observed for *Pa*DDAH interface area (3.33 Å), indicative of a loose interfacial packing of residues, this parameter was unable to distinguish clearly obligate from non-obligate dimers in the latest work from Thornton's laboratory (Ponstingl *et al.*, 2005).

In summary, these observations suggest that potentially none of *Pa*DDAH, AGAT or StrB1 enzymes requires self-association for biological activity. More generally, these considerations clearly demonstrate the general requirement to carefully assess the relevance of the self-associated state that is often observed in the 3D structures of proteins: by their nature such depictions of the protein result from interactions that occur at high concentrations (either in the crystal or in the NMR sample), and simple inspection of the structure does not reveal *per se* the thermodynamic characteristics of the association.

Individual residue contribution to PaDDAH interface ASA

To identify the residues that contribute to stabilizing interactions in the *Pa*DDAH homodimer, the predicted changes in solvent accessible surface area (ASA) upon dimerisation in a residue-by-residue fashion were examined (Table 3.2). The contribution of individual residues was investigated using the program NACCESS (Hubbard and Thornton, 1993). Values of solvent accessible surface area of each residue i in both the monomeric ($ASA_{i,M}$) and the homodimeric ($ASA_{i,D}$) species interface were derived and the change in surface area upon oligomerisation could be inferred for each individual residue (ΔASA_i) (Table 3.2). The same quantity could be expressed as an absolute value (Å^2) or as a fraction of $\Delta ASA_{i,M}$ ($fASA_i$). By this assessment a total of 23 residues undergo a finite change in ASA (i.e. $\Delta ASA_i > 0$ Å^2), of which 13 show a fractional change in ASA ($fASA_i$) greater than 0.6 and account for 567 Å^2 per protomer (66%) of the buried ASA. Plotting the values of $fASA_i$ on a 3D molecular surface of *Pa*DDAH clearly indicates the focal point of the intermolecular interaction which coincides with the exposed βC strand of module 1 composed mainly of hydrophobic residues (Figure 3.5a). From this representation, the entirety of the side chain ASA of Ile 49 and Leu 51 become concealed on the formation of the homodimer. However here, it is more helpful to consider the absolute residue-by-residue ΔASA values, which are shown in a colour-coded manner in Figure 3.4b. Two arginine residues (Arg 40 and Arg 98) are the largest contributors of side chain ASA to the interface, with a combined 25% of the total homodimer interface surface area between them. Both of these side chains form salt bridges with acidic residues on the symmetry-related protomer (Figure 3.2b,d).

The considerations derived from this analysis led us to limit the monomerisation mutagenesis strategy to residues Glu 33, Asn 36, Arg 40 and Asn 43 located in helix α1 and Arg 98 in helix α2. These residues have relatively high values of ΔASA_i and $fASA_i$. Although the hydrophobic core residues Ile 49, Leu 51 and Leu 52 located in β1C-strand are at the centre of the interface area and do contribute to the association of the homodimer with the three highest $fASA_i$ values, previous work performed in the laboratory indicated that point mutations aiming to disrupt these interactions led to the expression of insoluble protein (Plevin, 2003).

Table 3.2. Accessible surface area (ASA) statistics of residues at the homodimer interface of WT *Pa*DDAH from the Protein-Protein Interaction server[a]

Residue	$ASA_{i,M}$ (Å^2)[c]	$ASA_{i,D}$ (Å^2)[d]	ΔASA_i (Å^2)[e]	$fASA_i$
Lys 3	147.06	139.16	7.90	0.05
His 4	45.51	18.87	26.64	0.59
Thr 9	32.29	7.41	24.88	0.77
Leu 32	71.89	56.49	15.40	0.21
Glu 33	117.93	95.92	22.01	0.19
His 35	6.92	2.54	4.38	0.63
Asn 36	77.65	16.27	61.38	0.79
Ile 39	15.29	5.69	9.60	0.63
Arg 40	158.61	56.10	102.51	0.65
Gln 43	112.62	22.92	89.70	0.80
Asp 46	106.30	86.81	19.49	0.18
Val 47	19.08	2.92	16.16	0.85
Asp 48	81.54	19.49	62.05	0.76
Ile 49	46.66	0.97	45.69	0.98
Thr 50	30.37	2.15	28.22	0.93
Leu 51	86.62	1.69	84.93	0.98
Leu 52	22.77	5.78	16.99	0.75
Pro 53	89.25	37.26	51.99	0.58
Pro 54	64.78	49.37	15.41	0.24
Glu 94	133.50	119.34	14.16	0.11
Thr 95	5.86	5.12	0.74	0.13
Arg 98	148.36	39.79	108.57	0.73
Phe 99	56.23	27.65	28.58	0.51

[a]see Table 3.1; [b]ASA of residue *i* in monomer; [d]ASA of residue i in dimer; [e]difference between $ASA_{i,M}$ and $ASA_{i,D}$; [f]fractional ASA calculated where $fASA_i = (ASA_{i,M}-ASA_{i,D})/ASA_{i,M}$.

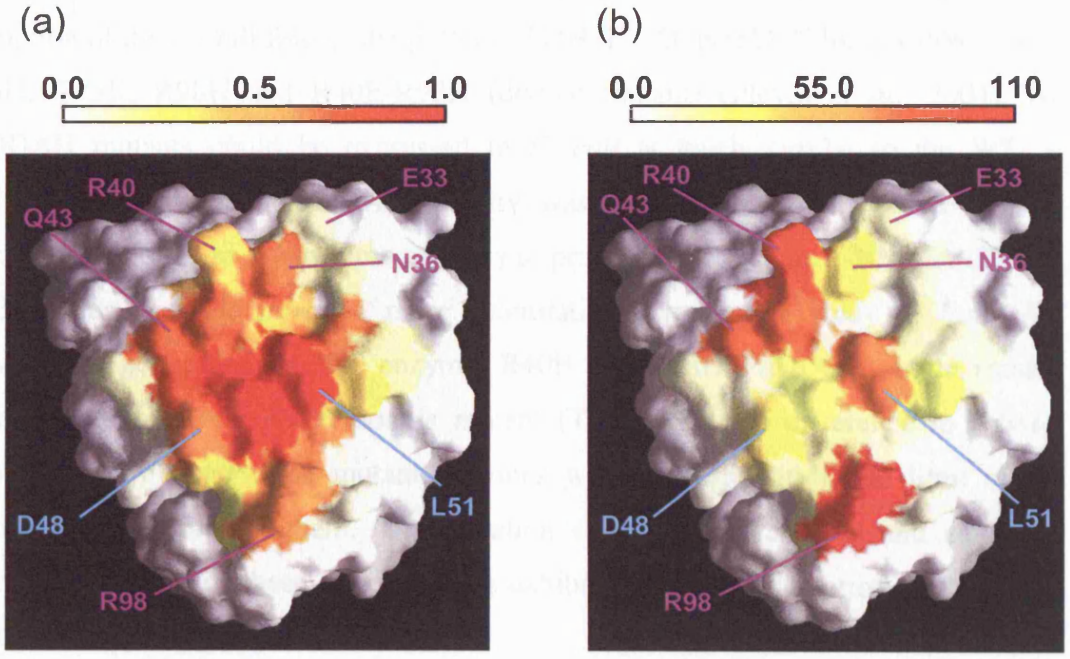

Figure 3.5. 3D structural analysis of the solvent-accessible surface area of WT *Pa*DDAH. (a) Molecular surface WT *Pa*DDAH coloured according to fractional ASA_i (see Section 3.3); (b) WT *Pa*DDAH with a molecular surface coloured according to absolute ΔASA_i ($Å^2$) (see Section 3.3). The orientation of the protomer is the same as the one presented in Figure 3.4a. Images were generated using GRASP (Nicholls *et al.*, 1991). Adapted from Plevin *et al.* (2004).

3.4 Biophysical characterisation of *Pa*DDAH interface mutants

Site directed mutagenesis of PaDDAH: assay of solubility and activity

On the basis of the ASA analysis, a total of ten *Pa*DDAH single residue mutants were designed in the lab and prepared using site-directed mutagenesis. All variants were selected to alter the physicochemical (charge and H-bonding potential) properties of the *Pa*DDAH protomer surface so as to undermine the self-association affinity without disruption of the overall fold of the protein: E33H, E33Q, N36D, N36H, N36W, R40E, Q43H, Q43R, R98H and R40E-R98H (double mutant) (Plevin *et al.*, 2004). All *Pa*DDAH mutants could be expressed in *E. coli* at levels similar to the WT. A qualitative assessment for DDAH activity was performed and all of the soluble mutants were found to retain some capacity to process the substrate L-NMMA (Plevin, 2003). Later on (*vide infra*), a more quantitative comparative study of *Pa*DDAH activity was performed on WT enzyme, R40E and R98H *Pa*DDAH single residue mutants and the R40E-R98H double mutant (Table 3.3). No difference in activity between the WT and these mutant enzymes was detected within the limit of the precision of the measurement. The retention of enzymatic activity and solubility indicates that none of these three mutants exhibit significant disruption to the active site of the enzyme.

Table 3.3. Comparison of WT and selected mutant *Pa*DDAH enzymatic activities

*Pa*DDAH isoform	Relative Activity [a]
WT	1.00 (0.04)
R40E	0.95 (0.12)
R98H	0.88 (0.13)
R40E-R98H	0.95 (0.04)

[a]Relative yield of L-NMMA to citrulline conversion assay (see Chapter 2, section 2.4), at a DDAH concentration of 4 μM. Values report activity relative to WT enzyme.

Analysis of the oligomeric state of PaDDAH mutants

The oligomeric state of each soluble *Pa*DDAH mutant was quickly assessed using analytical SEC at a single protein concentration (1 mg/ml; data not shown). From this initial characterisation, R40E, Q43R, R98H and R40E-R98H *Pa*DDAH displayed an

elution profile of single peak with an elution volume considerably greater than WT *Pa*DDAH (Plevin *et al.*, 2004). Interestingly, the N36W *Pa*DDAH mutant exhibited a lower elution volume than WT *Pa*DDAH with a more symetrical elution peak (data not shown). These proteins N36W, R40E, Q43R, R98H and R40E-R98H *Pa*DDAH were then selected for further assessment of the concentration dependence of their oligomeric state using a procedure similar to that applied to the WT enzyme. The elution profile for each of these proteins was obtained at six different sample dilutions over the range 5 mg/mL to 50 μg/mL. In each case the elution volume of the major band was converted into an apparent molecular weight (MW_{app}) value by reference to the standard calibration plot (Figure 3.6a). At difference to what was observed for WT *Pa*DDAH, the apparent molecular weights (MW_{app}) obtained for R40E, Q43R and R98H *Pa*DDAH clearly decrease to values that cluster between 27-33 kD, a result that suggests to a more monomeric behaviour in solution but that is still dependent on protein concentration. Intriguingly the mutant N36W *Pa*DDAH showed a constant MW_{app} at around 51 kD as it appears that this protein is essentially dimeric. Moreover, the double mutant R40E-R98H *Pa*DDAH showed no concentration dependence of MW_{app} over the concentration range sampled, consistent with an invariant MW_{app} of 29 kD (Figure 3.6a).

In order to quantify further the self-association of WT, N36W, Q43R, R40E, R98H and R40E-R98H *Pa*DDAH proteins, sedimentation equilibrium analytical ultracentrifugation (SE-AUC) experiments were also performed. The experiments were designed to sample the dependence of the sedimentation behaviour upon the centrifugation speed and protein concentration. Several protein concentrations for each individual protein were sampled at three different rotor speeds (see Chapter 2, Section 2.4). However, to streamline the analysis, the interference datasets collected at 20,000 revolutions per minute (rpm) were initially fit to a single ideal species model (see Equation 2.1, Chapter 2, Section 2.4). The resulting weight-averaged molecular masses ($M_{w,app}$) for each protein are shown as a function of protein concentration in Figure 3.6b. Three types of hydrodynamic behaviour are readily apparent. First, WT and N36W *Pa*DDAH yield $M_{w,app}$ values around 55-57 kD. Second, the $M_{w,app}$ of Q43R, R40E and R98H *Pa*DDAH varied from 36 kD to 46 kD with increasing protein concentration. Finally, the double mutant R40E-R98H *Pa*DDAH presented a consistent $M_{w,app}$ around 27 kD over the concentration range. In qualitative terms these

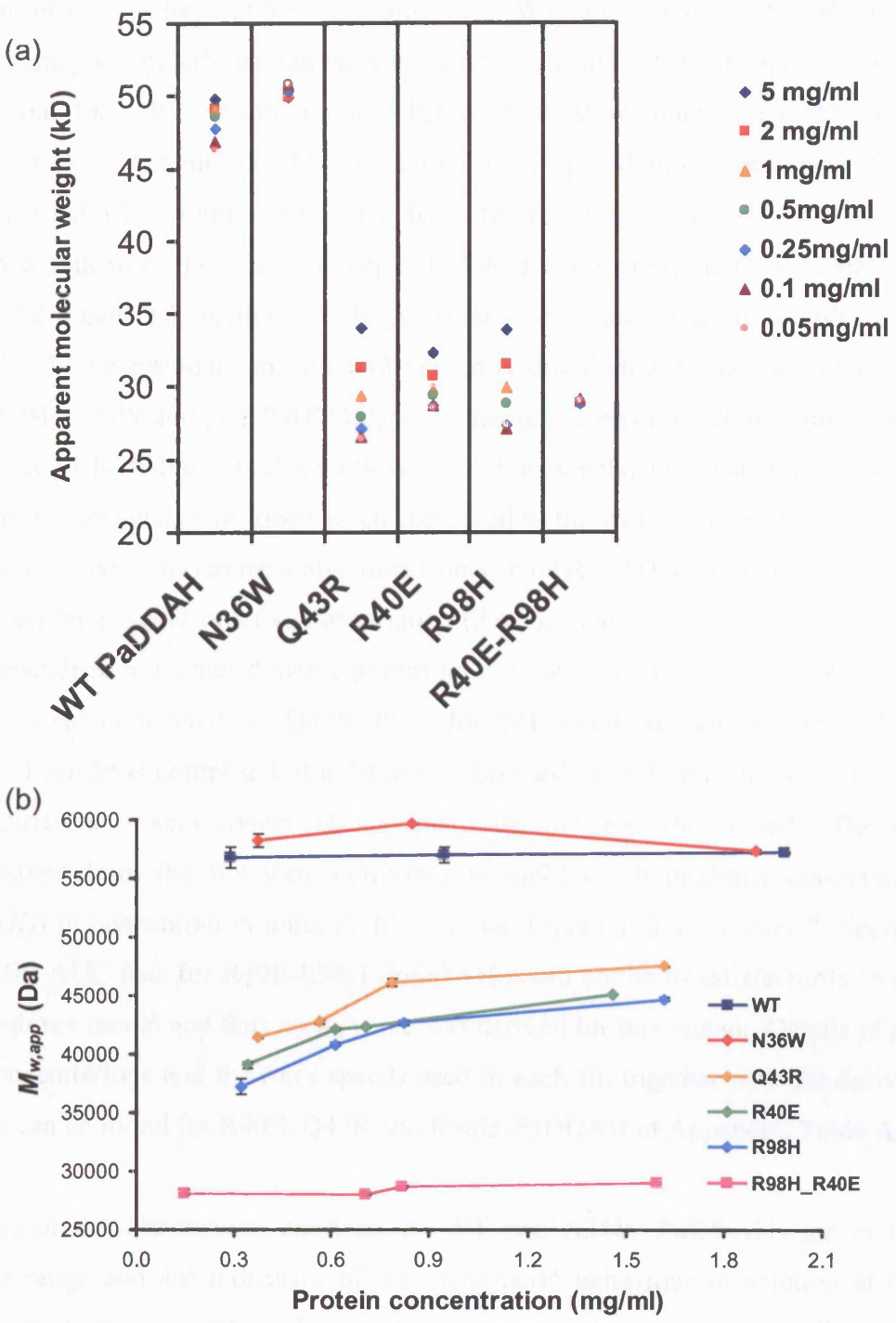

Figure 3.6. (a) Apparent molecular weight of WT and mutant *Pa*DDAHs at varying loading concentrations indicated in the colour-coded legend. (b) Sedimentation equilibrium analytical ultracentrifugation analysis of WT *Pa*DDAH and *Pa*DDAH mutants. Apparent weight-averaged molecular masses ($M_{w,app}$) of WT *Pa*DDAH, N36W, Q43R, R40E, R98H and R40E/R98H *Pa*DDAH as a function of cell loading concentration. The plotted $M_{w,app}$ values were obtained by averaging the results from the individual curve fits to a single ideal species model of three interference subsets collected at either 19,000 rpm (WT *Pa*DDAH) or 20,000 rpm (all mutant DDAHs).

results are broadly consistent with the SEC data: WT and N36W *Pa*DDAH show a strong tendency to homodimerization, whereas R40E-R98H *Pa*DDAH appears to lack any potential for self-association; the other mutants show intermediate behaviour consistent with a dynamic equilibrium between dimer and monomer forms. In an attempt to establish a more quantitative basis for these observations, the SE-AUC interference data were then fit to an explicit monomer-dimer equilibrium model (see Equation 2.2, Chapter 2, Section 2.4). Representative curves of these fits are illustrated in Figure 3.7. The residuals are low and randomly distributed. In the case of R40E, Q43R, R98H, N36W and WT *Pa*DDAH, when the monomer-dimer equilibrium model fit was applied both the visual quality of the fit was enhanced, and the statistical variance of the residuals was lower when compared to the values obtained for a single ideal species model, monomer-only, dimer-only models (Table 3.4) or for fits to models describing higher-order self-association (data not shown).

Four independent monomer-dimer equilibrium fits were performed for R40E and R98H, six were performed for Q43R, three for WT *Pa*DDAH and two for N36W *Pa*DDAH. Each fit is composed of a dataset of nine sedimentation equilibrium curves at three different protein concentrations and three different rotor speeds. The K_{a2} values derived from the fits were converted to yield the homodimer dissociation constant (K_d) in concentration units (Table 3.4; see Equation 2.3, Chapter 2, Section 2.4). The SE-AUC data for R40E-R98H *Pa*DDAH could not be fit satisfactorily to the monomer-dimer model and thus no K_d value was derived for this mutant. Details of the protein concentrations and the rotor speeds used in each fit, together with the derived K_{a2} values can be found for R40E, Q43R and R98H *Pa*DDAH in Appendix, Table A.I.

The values of the dissociation constants of WT and N36W *Pa*DDAHs are in the nanomolar range and are indicative of a homodimeric behaviour in solution at the protein concentrations sampled. Conversely the single point mutations Q43R, R40E and R98H substantially disrupted the association between protomers; all have K_d values in the micromolar range, on average 100-fold higher than the equilibrium dissociation constant for the WT protein. More importantly, the SE-AUC data clearly show that R40E-R98H *Pa*DDAH is a variant of *Pa*DDAH that lacks evidence for a monomer-dimer equilibrium under the conditions of these experiments.

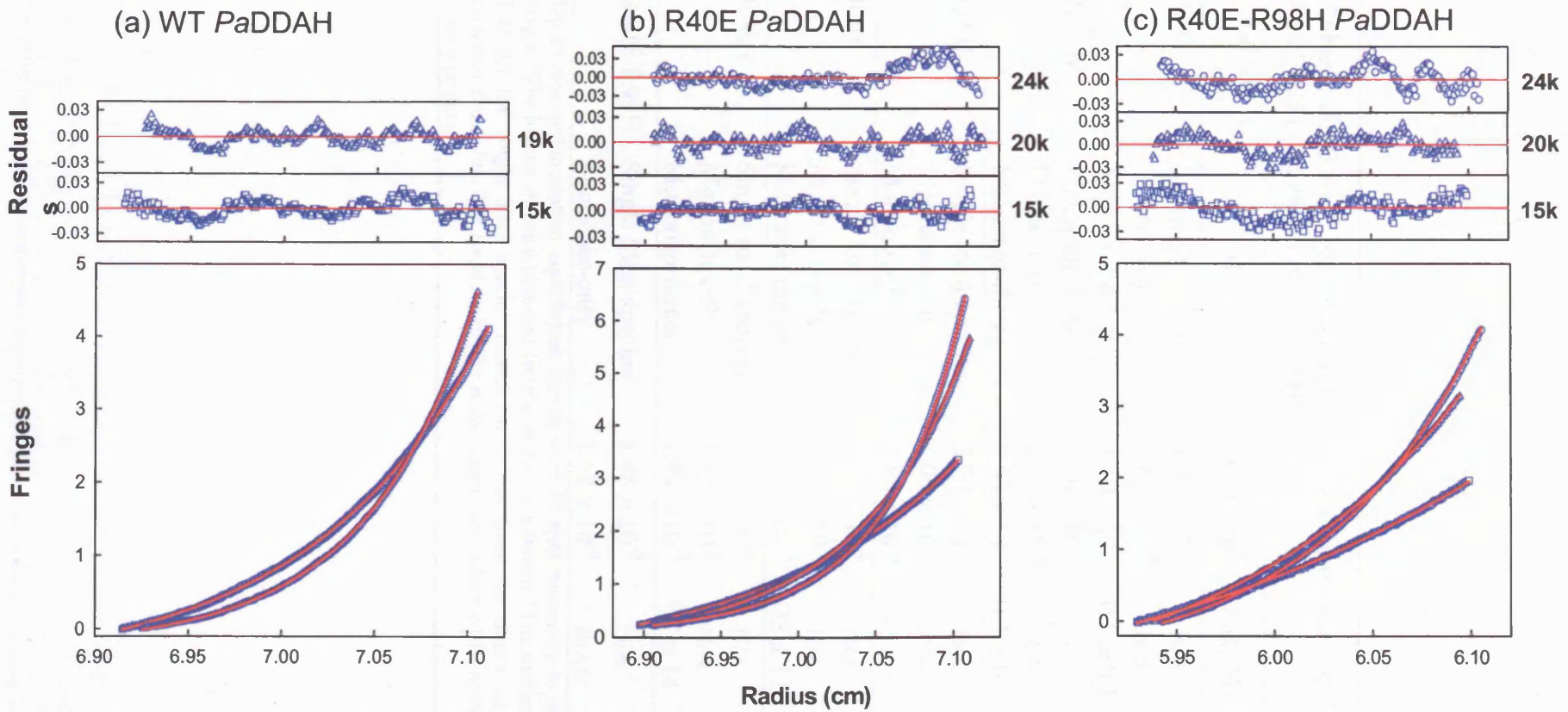

Figure 3.7. Representative examples of the sedimentation equilibrium curve fits for (a) WT *Pa*DDAH (0.94 mg/mL), (b) R40E *Pa*DDAH (0.71 mg/mL), and (c) R40E-R98H *Pa*DDAH (0.82 mg/mL). In (a) and (b) the interference data (blue symbols, lower panels) are fitted (continuous red line, lower panels) to a monomer-dimer equilibrium; in (c) the fit is to a single ideal species model. Each symbol corresponds to a different rotor speed (in rpm) as indicated on the right hand side of the fit residuals (plotted on an expanded scale in the corresponding upper panels). In order to derive a single Ka_2 value for R40E PaDDAH monomer-dimer equilibrium, two other concentrations at three different speeds were used simultaneously (see Appendix I, Table A.I).

Table 3.4. Fitting of equilibrium sedimentation data of WT *Pa*DDAH and *Pa*DDAH mutants to various global models

*Pa*DDAH isoform	Model[a]	Variance[b]	K_d (μM)	Average molecular mass (kD)[c]
WT	Single ideal species	1.12×10^{-4}	N/A	57.0
	Dimer-only	1.84×10^{-4}	N/A	58.4
	Monomer-dimer	1.10×10^{-4}	0.5 ± 0.1	29.2
N36W	Single ideal species	1.66×10^{-4}	N/A	57.0
	Dimer-only	2.54×10^{-4}	N/A	58.6
	Monomer-dimer	1.62×10^{-4}	0.3 ± 0.1	29.3
Q43R	Single ideal species	2.70×10^{-4}	N/A	46.3
	Monomer-only	1.00×10^{-2}	N/A	29.2
	Monomer-dimer	1.55×10^{-4}	45 ± 5	29.2
R40E	Single ideal species	3.07×10^{-4}	N/A	42.6
	Monomer-only	1.19×10^{-2}	N/A	29.2
	Monomer-dimer	8.53×10^{-5}	73 ± 11	29.2
R98H	Single ideal species	2.59×10^{-4}	N/A	41.6
	Monomer-only	3.45×10^{-3}	N/A	29.2
	Monomer-dimer	1.99×10^{-4}	87 ± 14	29.2
R40E-R98H	Single ideal species	1.48×10^{-4}	N/A	28.6
	Monomer-only	1.68×10^{-4}	N/A	29.2

[a]Up to nine sedimentation equilibrium curves were fit simultaneously to one global model using Origin. [b]The lowest variance obtained from a single fit is shown. [c]The molecular mass is allowed to float for the single ideal species model fit, it was fixed for dimer-only, monomer-only and monomer-dimer fits. The fixed molecular mass values were obtained from the amino acid sequence of each protein.

Thermodynamic considerations

To take these findings further, the determination of the dissociation constants of the various *Pa*DDAH mutant homodimers is a prerequisite to the thermodynamic study of the contribution of each residue side chain on the free energy of association. The common assumption is that the energy of protein-protein associations is strongly related to the buried hydrophobic surface area (Chothia and Janin, 1975, Jones and Thornton, 1996). However, when considering individual residue energy contributions instead of the interface surface as a whole, Wells and colleagues have demonstrated through alanine scanning mutagenesis that a single side chain can contribute a large fraction of the binding free energy (Clackson and Wells, 1995). Moreover, Bogan and Thorn compiled a dataset of interface residues mutated to alanine in various protein complexes (Bogan and Thorn, 1998). Their analysis has shown that there is little correlation between side chain buried surface area and free energy of association. In fact, in agreement with Wells' studies not only do certain side chains contribute disproportionately more to the bulk of the energy of association but a bias exists for tryptophan, arginine and tyrosine residues to appear in more than 10% of the cases in what the authors have termed "hot spots". Following their definition, a hot spot is a residue that when mutated to alanine gives rise to a distinct increase of more than 2 kcal/mol in the absolute binding free energy. The Bogan and Thorn findings imply that residues located in a hot spot do not share similar physico-chemical properties such as hydrophobicity or charged state. Rather, amino acids capable of forming multiple types of favourable interactions are preferred in interface environments. Tryptophan for example can contribute to the binding energy through aromatic interactions, a hydrogen bonding donor and a large hydrophobic contact. Likewise tyrosine offers a hydrophobic surface, aromatic interactions and the potential to form a hydrogen bond with the ring hydroxyl group. As for arginine residues, they can form up to five hydrogen bonds and a salt bridge with the guanidinium moiety. In addition, the π-electron delocalisation of the guanidinium group has a pseudo-aromatic character and the three methylene carbon atoms have a hydrophobic character.

In our case, the mutations of residues Arg 40 and Arg 98 contributed to a decrease in free energy of homodimer dissociation $\Delta(\Delta G_d)$ of 3 kcal/mol, and that of Gln 43 contributed 2.6 kcal/mol relative to the WT protein (Table 3.5). Although these mutations are not to alanine residues, Arg 40, Gln 43 and Arg 98 can equally be defined as binding or interfacial hot spots.

Table 3.5. Corresponding free energy (ΔG_d) of binding and individual free energy contributions to binding $\Delta(\Delta G_d)$ for WT and mutant *Pa*DDAH

*Pa*DDAH isoform	K_d (μM)	ΔG_d (kcal mol^{-1})[a]	$\Delta(\Delta G_d)$ (kcal mol^{-1})[b]
WT	0.5 ± 0.1	8.4 ± 1.7	N/A
N36W	0.3 ± 0.1	8.8 ± 3.4	-0.4 ± 0.1
Q43R	45 ± 5	5.8 ± 0.6	2.6 ± 0.4
R40E	73 ± 11	5.5 ± 0.8	2.9 ± 0.3
R98H	87 ± 14	5.4 ± 0.9	3.0 ± 0.4

[a]The free energy of dissociation ΔG_d was obtained with the following equation: $\Delta G_d = -RT\ln K_d$, where R is the gas constant (1.987 cal·K^{-1}·mol^{-1}); T is the temperature in Kelvin (293K). $\Delta(\Delta G_d) = \Delta G_{d,WT} - \Delta G_{d,Mutant}$

However, in the case of the R40E mutation, the results are not informative as to whether the destabilisation effect originates from the disruption of a salt-bridge or from the electrostatic repulsion created by the insertion of a negatively charged residue in position 40, in close proximity to Glu 33 on the opposing helix-α1 (Figure 3.2d). In fact, it is likely that the destabilisation of the protein-protein interactions caused by these mutations arises from numerous sources. In our mutagenesis strategy the mechanism involved is most likely the loss of electrostatic pairing, with the disruption of two pairs of salt-bridges (Arg 40-Glu 33 and Arg 98-Asp 48). However, it is possible that other mechanisms, not easily recognised, could account to the overall energetic contribution of destabilisation of the dimer. These include a discrete change in local conformation due to the accommodation of a novel residue, increased entropy of dissociated monomers, or a change of the ionisation state of charged residues at the mutated interface (DeLano, 2002).

On a practical note, to our surprise, N36W *Pa*DDAH was found to have a lower K_d than WT *Pa*DDAH and its hydrodynamic properties point to an exclusive dimeric behaviour in solution. This mutation was intended to add a bulkier side chain in the interface and hence obstruct the proximity of the protomers that would hinder the association. As mentioned above, tryptophan residues are prone to form several interactions in an interface environment. Its presence in position 36 of helix $\alpha1_A$, in close proximity to the reciprocal Trp residue in the opposing helix-$\alpha1_B$ perhaps creates additional hydrophobic interactions, sufficient to account for the small increase in free

energy of binding (Figure 3.2d, Table 3.5) and to disrupt the apparent monomer-dimer equilibrium as demonstrated by the SEC and SE-AUC results.

Initial analysis of monomeric PaDDAH by NMR spectroscopy

2D [^{1}H,^{15}N]-HSQC NMR spectra were recorded of [^{15}N]-isotope labelled samples of WT *Pa*DDAH and selected mutant forms. WT *Pa*DDAH yields a reasonably well-dispersed spectrum with *ca.* 160 resolvable cross peaks (Figure 3.8a). The aim was to obtain spectra with a complete tally (241) of cross peaks (*i.e.* 254 less 13 prolines). However, the single *Pa*DDAH mutants failed to improve the quality of the NMR spectra (data not shown), as no additional cross peaks were observed compared to the number normally obtained for the WT protein (Plevin, 2003). The derived dissociation constants for these mutants are in the micromolar range which helped us in the understanding that from an NMR perspective, they are still essentially dimeric at the concentrations required for NMR studies (Figure 3.9). However the HSQC spectrum of [^{15}N]-labelled R40E-R98H *Pa*DDAH (Figure 3.8b) shows a marked enhancement in the quality compared to the WT protein. Simple inspection reveals that the individual NH cross peak line widths are significantly reduced and are accompanied by an apparent increase in, and uniformity of, signal-to-noise (S/N) ratio. The central region of the spectrum, which is highly overlapped in the WT spectrum, shows a clear improvement in signal resolution, leading to a provisional tally of as many as 241 backbone NH cross peaks, including two cross peaks with H^N chemical shifts at around 4.8 ppm close to the noise band derived from the water signal that had not previously been properly discernible in the spectrum of the WT protein.

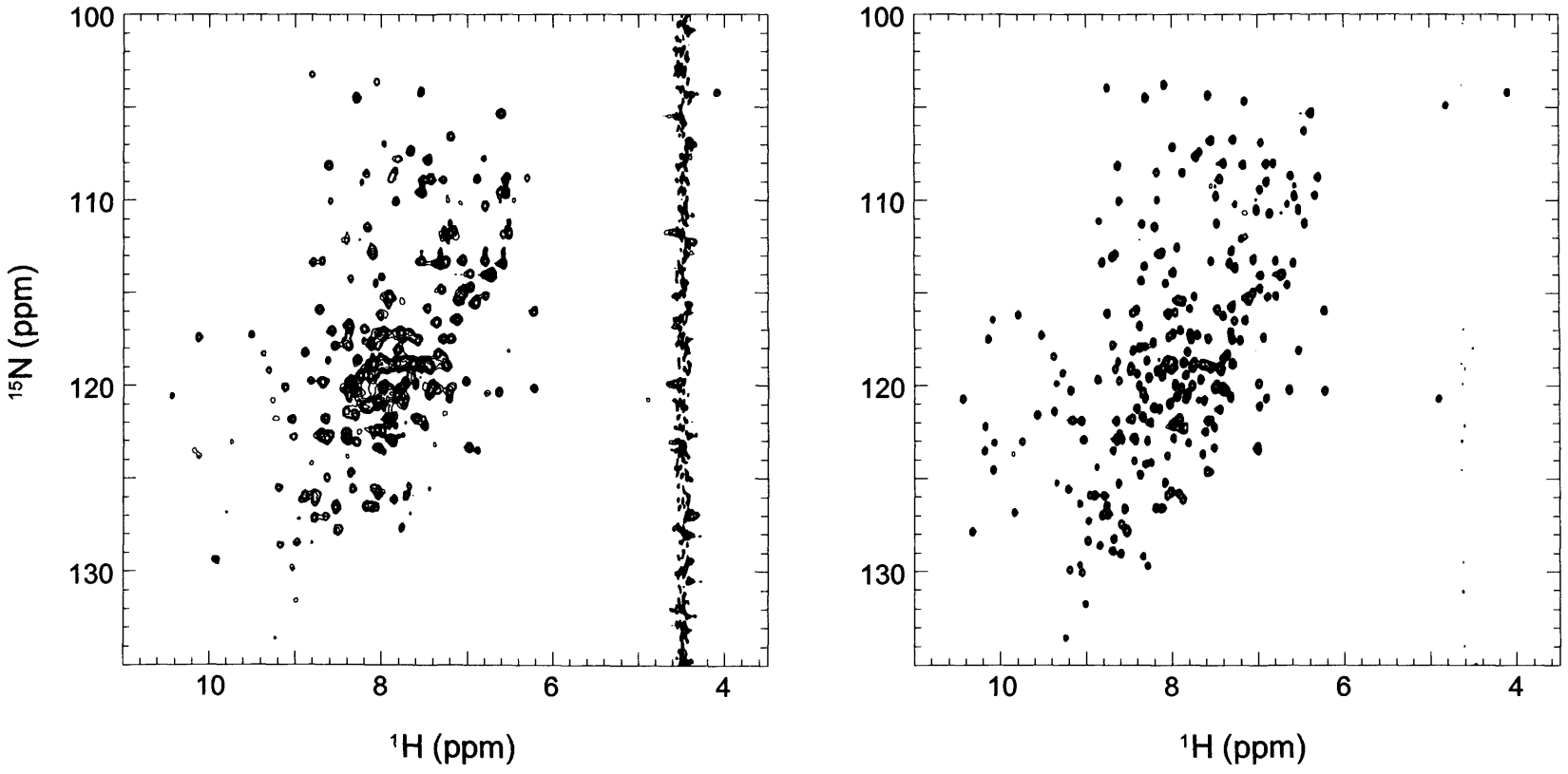

Figure 3.8. Comparison of [^{1}H,^{15}N]-HSQC spectra recorded of [^{15}N]-labelled WT *Pa*DDAH (a) and R40E-R98H *Pa*DDAH (b). Each experiment was recorded under identical experimental parameters at a protein concentration of 1.0 mM at a proton frequency of 600 MHz. Each spectrum shows a region of a data matrix that consisted of 400* t_1 and 2048* t_2 data points (where n* refers to complex points) with acquisition times of 90 ms (t_1) and 256 ms (t_2). A total of 32 transients were recorded per t_1 increment, with a repetition rate of 1.7 s and total experiment duration of *ca.* 6 h. The spectra are not plotted at the same contour levels: the overall lower signal-to-noise ratio of the WT spectrum (a) means that at the lower contour level required to make the majority of cross peaks observable in the plot, the residual solvent suppression noise at ^{1}H ~ 4.6 ppm is more plainly visible.

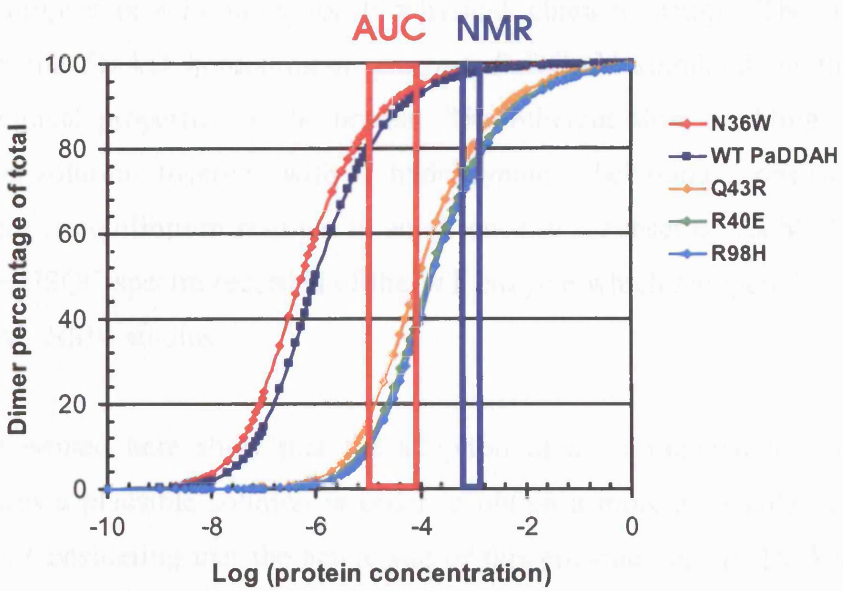

Figure 3.9. Simulated curves of dimer percentage as a function of total protein concentration obtained from the derived K_ds of each *Pa*DDAH protein (see Chapter 2, Section 2.4). The red and blue rectangles highlight the protein concentration range needed in AUC and NMR experiments respectively.

3.5 Discussion

The engineering of monomeric proteins from higher order oligomers is not a novel idea. Generally, these studies are based on known three-dimensional protein structures and aim to explore oligomer association and its relative importance for biological activity (*e.g.* Borchert *et al.*, 1994, Rajarathnam *et al.*, 1994, Sakurai and Goto, 2002, Cox *et al.*, 2003). An additional premise for the adoption of such a strategy is that it can offer simpler protein units for biophysical characterisation. The attempts to characterise the 58 kD homodimeric enzyme *Pa*DDAH stumbled on the intrinsic physico-chemical properties of the protein. The inherent slow tumbling of a larger molecule in solution together with a hydrodynamic behaviour consistent with a monomer-dimer equilibrium resulted in an absence of a subset of visible H^N-N cross peaks in the HSQC spectra recorded of the WT enzyme which hampered our ability to pursue further NMR studies.

The data presented here show that the adoption of a monomerisation strategy for *Pa*DDAH was a plausible solution in order to obtain a more amenable molecule for NMR study. Considering that the active site of this enzyme lies *ca.* 15 Å away from the interface region, it was likely that the enzymatic activity of a monomeric variant would not be at risk if residues at the interface were targeted for mutagenesis. In support of this scenario, we showed, based on the available crystal structure, that the interface surface area of the homodimer is small (860 $Å^2$) when compared to other stable oligomers, it is not compact as demonstrated by the high Gap Index value (3.3 Å), and it has a high content of hydrophilic residues reflecting a composition similar to the solvent exposed protein surface (Table 3.1). Since in many systems the biological role of oligomerisation is not clear, we postulated that the dimeric nature of *Pa*DDAH was the result of the high concentrations needed for crystallographic and NMR studies.

Subsequent analysis of the solvent accessible surface area of the individual residues involved in the association led us to target Glu 33, Asn 36, Arg 40, Gln 43 and Arg 98 for site-directed mutagenesis. These residues demonstrate high absolute changes in both ASA and in $fASA_i$ upon dimerisation. In particular, the mutation R40E aimed to disrupt a pair-wise salt-bridge formed with Glu 33 in the opposing protomer. As for Q43R mutation, the additional charge introduced by the guanidino group of the

arginine residue is predicted to increase the electrostatic repulsion between position Arg 43 and the opposing Arg 40 across $\alpha 1_{A,B}$ helices. The mutation of Arg 98 to His located in helix-α2 was intended to disrupt a salt-bridge interaction with the carboxylate group of Asp 48 in strand β1C. Under conditions in which the introduced His is in the deprotonated state, the electrostatic interaction with Asp 48 is unlikely to occur.

All three single residue mutants showed some degree of deviation from the hydrodynamic behaviour of the WT enzyme when characterised by analytical SEC and SE-AUC. The latter technique revealed a shift in the weight-averaged apparent molecular mass from 57 kD as determined for WT *Pa*DDAH to an average of 43 kD. The quantitative estimates of the dimer dissociation constants for these proteins showed a 100-fold decrease in their homodimer association which, nevertheless, was unfortunately insufficient to improve the quality of the NMR spectra (Plevin, 2003); these enzymes are essentially 80% dimeric at typical NMR concentrations ($\approx$ 0.5 mM).

With regard to N36W *Pa*DDAH, its apparent molecular mass remained unaffected by changes in protein concentration and the increase in free energy of association strongly suggests a shift of the oligomeric equilibrium towards a consistent dimeric behaviour in solution. N36W *Pa*DDAH is a more stable homodimer and this fact led us to characterise this mutant by NMR spectroscopy with the notion that the slight increase in the association strength would lessen the effect of any chemical exchange induced line broadening due to monomer-dimer equilibrium at the interface. However, the 2D HSQC spectrum of [^{15}N]-labelled N36W *Pa*DDAH had similar characteristics to that of WT *Pa*DDAH and several resonances were still missing.

In summary, it appears that the combination of two single point mutations efficiently disrupt the association of *Pa*DDAH monomers. The R40E-R98H *Pa*DDAH double mutant provides an excellent platform to study the structure-function characteristics of *Pa*DDAH as a stable active monomer species. The overall excellent stability and hydrodynamic properties of [^{15}N]-labelled R40E-R98H *Pa*DDAH, clearly indicate that it should be within the scope of triple resonance NMR spectroscopic methods to obtain a full resonance assignment that will provide the basis for the precise characterisation

of residue-by-residue dynamics properties and the substrate and/or inhibitor ligand binding effects.

The following chapter will describe the full assignment of the spectrum of ^{15}N-labelled R40E-R98H *Pa*DDAH which helped in the understanding of the limited utility of the NMR data for the WT protein. Despite its well dispersed 2D [^{1}H-^{15}N]-HSQC spectrum, the poor quality of the spectrum of WT *Pa*DDAH was not unexpected, considering the relatively large size of the homodimeric WT protein (MW_{app} ~ 58 kD). Any globular molecule of this size would be predicted to exhibit a relatively long rotational correlation time leading to rapid decay of transverse magnetization and broad NMR features. However, this slow tumbling characteristic alone may not be sufficient to account for the apparent absence of a significant fraction of the expected cross peaks. Possible explanations for this phenomenon will be discussed in Chapter 4.

Chapter 4

*Pa*DDAH backbone resonance assignments

Abstract

Recently developed NMR techniques such as transverse relaxation optimised spectroscopy (TROSY) applied in combination with the routine isotopic enrichment of proteins have permitted the detailed NMR characterisation of systems of up to 100 kD. These advances should have enabled the study the 58 kD WT *Pa*DDAH enzyme but the efforts were hampered by intrinsic physico-chemical properties of the homodimer which lead to a substantial loss of the expected resonances. This chapter describes the application of multidimentional NMR methodology to obtain the sequence specific backbone resonance assignments of the newly engineered 29 kD monomeric variant of the *Pa*DDAH enzyme. The smaller size of the protein compared to the wild-type protein significantly enhanced the resolution of the central region of the 2D spectra recorded and contributed to the overall increase of the signal-to-noise ratio. The improvements in the quality of the spectra enabled the full assignment of the backbone ^{1}H, ^{13}C, and ^{15}N resonances of [^{2}H,^{13}C,^{15}N]-labelled R40E-R98H *Pa*DDAH with a series of triple resonance experiments. The deuteration of the protein proved crucial to enhance the efficiency of the magnetisation transfer steps in triple resonance experiments. Furthermore the ^{13}C chemical shift data were analysed as a means of identifying the locations of the secondary structure elements of *Pa*DDAH and a good agreement was observed between the NMR data and the available crystal structure of WT *Pa*DDAH. This level of NMR characterisation allowed for the identification of residues for which the $^{1}H^{N}$, ^{15}N resonances were absent in the 2D spectra of the WT *Pa*DDAH. Interestingly, this assessment clearly shows that cross peaks that are lost in the fully protonated sample correspond to residues located at the interface region of the homodimer and to residues located in the surrounding loops, such as all the residues involved in catalysis.

4.1 The study of large proteins by NMR

Isotopic enrichment of proteins

The characterisation of large proteins by NMR spectroscopy has benefited over the last two decades from significant improvements in experimental protocols and technical developments. The isotopic enrichment of proteins with ^{13}C and ^{15}N isotopes is now employed in a routine manner in NMR laboratories. Although seemingly adding a layer of complexity, backbone and side chain $^{1}H^{N}$, ^{13}C and ^{15}N spins can be correlated by transferring magnetisation from one nucleus to the other through one- or two-bond scalar couplings or through space, greatly reducing ^{1}H signal overlap. For many proteins with a molecular weight around 20 kD, the transfer of magnetisation between nuclei via scalar couplings is faster than the decay of magnetisation from relaxation so that high sensitivity spectra can be recorded. Hence, several triple-resonance pulse sequences in three and four dimensions (3D, 4D) have been developed which allow the full assignment of proton, nitrogen and carbon backbone and side chain chemical shifts of proteins with molecular weights of up to approximately 20-25 kD (Clore and Gronenborn, 1998, Sattler *et al.*, 1999).

Despite the utility of triple-resonance multi-dimensional NMR spectroscopy in structural studies, for systems larger than 25 kD, long molecular correlation times result in rapid relaxation rates of all nuclei, specially of the proton attached to ^{13}C nuclei, which attenuates signal intensity and decreases spectral resolution (Gardner and Kay, 1998). In the case of the ^{13}C-^{1}H and ^{15}N-^{1}H spin-pairs in a protein, the major contributor to relaxation of carbon and nitrogen magnetisation derives from the dipolar interaction between the two bonded atoms and in part to chemical shift anisotropy (CSA). The rate of dipolar-mediated relaxation is dependent on the product of the gyromagnetic ratios γ of the bonded nuclei. From a simplistic point of view, the larger this product, the greater the rate of dipolar relaxation. In order to study larger proteins, one method to circumvent this phenomenon has been to substitute non-exchangeable protons with deuterium nuclei. Deuterium has a significantly lower gyromagnetic ratio when compared to proton ($\gamma[^{1}H]/\gamma[^{2}H] \approx 6.5$). When ^{2}H replaces ^{1}H in carbon-bearing positions in proteins, not only do the contributions from proton-proton dipolar relaxation and ^{1}H-^{1}H scalar couplings to proton line widths decrease, but the dipolar interaction between a ^{13}C-^{2}H pair is attenuated. In fact, the likely contribution of the dipolar relaxation of ^{2}H to the ^{13}C line width for a protein undergoing isotropic

tumbling with a rotational correlation time of 15 ns is 1.5 Hz instead of 20 Hz in the case of a ^{13}C-^{1}H spin-pair, which consequently improves the sensitivity of the triple resonance experiments based on recording the NH group (Grzesiek *et al.*, 1993). Similarly, deuterating all non-exchangeable sites in proteins leads to a decrease in amide proton transverse relaxation ($^{1}H^{N}$ T_2) by a factor of two, as evidenced in the studies of villin 14T (Markus *et al.*, 1994) significantly improving the sensitivity and resolution of the acquired spectra (Markus *et al.*, 1994). These results have been further complemented by the studies of the 37 kD trp repressor/DNA complex, for which a suite of triple-resonance experiments for the assignments of the backbone resonances of [^{2}H,^{13}C,^{15}N]-labelled proteins was developed (Yamazaki *et al.*, 1994a,b). The level of protein deuteration used in Yamazaki *et al.* studies was 70% which is proportional to the level of D_2O/H_2O used to culture the bacteria containing the expression vector. For their purposes, which were to pursue structural studies, the preparation of a fully deuterated sample combined with reprotonation of backbone amide positions in H_2O-based buffers would ultimately preclude its use for experiments which require the magnetisation to originate or be detected in carbon-bearing protons. The depletion of protons in a perdeuterated protein sample strongly impairs nuclear Overhauser enhancement (NOE) based experiments, which rely on the through space correlation of protons. For instance, perdeuteration of the PLCC SH2 domain has been shown to lead to a five-fold decrease in the number of protons present (Gardner *et al.*, 1997), which greatly reduced the number of observable distance restraints between proton methyl groups. However alternative isotope labelling strategies have been elegantly applied to selectively incorporate ^{1}H into non-exchangeable sites of highly perdeuterated proteins. To this end, the supplementation of specific protonated amino acid precursors in D_2O-based growth media can be adopted to incorporate in proteins protonated methyl groups of alanine, valine, leucine and isoleucine (γ_2 only). These strategies have been shown to balance effectively both the chemical shift assignments and the NOE information necessary to determine protein 3D solution structures (Rosen *et al.*, 1996, Gardner *et al.*, 1997, Goto *et al.*, 1999). To date, the different protein sites that can be selectively isotopically labelled vary according to the type of precursor added to the growth media (Tugarinov *et al.*, 2004, Ollerenshaw *et al.*, 2005). It has become increasingly clear that the optimal labelling pattern of a protein depends upon the level of NMR characterisation required.

Transverse relaxation optimised spectroscopy (TROSY)

The recently developed principles of transverse relaxation optimised spectroscopy (TROSY; Pervushin *et al.*, 1997) significantly extend the size of proteins tractable by NMR to molecular masses beyond 100 kD. The main advantage of TROSY is the suppression of transverse relaxation during the magnetisation transfer periods and the evolution periods that leads to signal loss and line broadening of the NMR signal. Although the relaxation rates of amide protons in proteins in the 20-25 kD range have been successfully reduced by the partial or complete replacement of protons in non-labile positions by deuterons, TROSY further suppresses transverse relaxation in $^{1}H^{N}$-^{15}N moieties by constructive use of interference between dipole–dipole coupling and chemical shift anisotropy (CSA). In conventional [^{1}H,^{15}N]-heteronuclear spectroscopy, the doublet that arises from the scalar coupling between the two nuclei is seen in 2D experiments as a four-fold multiplet structure. Usually, the four-line component is effectively turned into a single centrally located line through broadband ^{15}N-decoupling during the evolution and acquisition periods which selectively removes the spin-spin splittings in the final spectrum. The underlying technical principle of TROSY is that the NH signal is retained in the multiplet state. Since the individual multiplet components have different transverse relaxation rates due to cross-correlation and hence different line widths, TROSY is engineered to select the narrowest, most slowly relaxing line of each multiplet (Riek *et al.*, 2000). For proteins with a long correlation time, the resulting sharp component improves spectral resolution and overall sensitivity. However, in order for the application of TROSY to be most effective, the contributions to relaxation from homonuclear proton-proton couplings have to be considered. In fact, the relaxation pathway of ^{1}H-^{15}N spin-pairs due to the dipolar couplings with remote protons cannot be influenced by TROSY. Therefore, an optimal TROSY effect is achieved with the use of uniformly deuterated proteins which significantly limits the coupling between adjacent amide and carbon-bearing protons.

In this regard, the application of TROSY-based pulsed sequences in conjunction with the deuteration of large proteins led to the complete backbone assignments of a 110 kD homooctameric protein, 7,8-dihydroneopterin aldolase (Salzmann *et al.*, 2000) and that of the human tumor suppressor protein p53 of 67 kD (Mulder *et al.*, 2000). More recently, the 3D solution structures of two eight stranded β-barrel membrane proteins were solved, OmpX (Fernandez *et al.*, 2004) and *Pag*P (Hwang *et al.*, 2002) dissolved in lipid detergents with effective molecular mass in the 50–60 kDa range. In addition, the 3D structure of the 80 kD bacterial malate synthase G was successfully solved by

the exploitation of TROSY-based approaches in combination with valine, leucine and isoleucine proton methyl labelling (Tugarinov *et al.*, 2005).

4.2 Assessing the need for TROSY methodology and deuteration in the study of dimeric and monomeric *Pa*DDAH

The study of WT PaDDAH by heteronuclear NMR

Previous work done in our laboratory demonstrated the significant improvement that the combination of TROSY-based methods and deuteration brought to the study of the 58 kD homodimeric WT *Pa*DDAH (Plevin, 2003). Examples of [^{1}H,^{15}N]-heteronuclear single quantum coherence (HSQC) and [^{1}H,^{15}N]-TROSY spectra recorded of [^{1}H,^{15}N]- and [^{2}H,^{15}N]-labelled WT *Pa*DDAH are shown in Figure 4.1 and clearly indicate an increased spectral resolution and signal-to-noise (S/N) ratio when [^{1}H,^{15}N]-TROSY is applied to a perdeuterated *Pa*DDAH sample (Figure 4.1c; Plevin, 2003). For our purposes, which were to obtain the sequence-specific resonance assignments of *Pa*DDAH, the TROSY-based triple resonance experiments (Yang and Kay, 1999) applied to a fully [^{2}H,^{13}C,^{15}N]-triple labelled protein sample were thought likely to provide the highest sensitivity and resolution to study the 58 kD protein. However, the lack of a subset of expected cross peaks in [^{1}H,^{15}N]-correlated spectra of both fully protonated and perdeuterated protein samples together with the relatively low number of HN $\rightarrow$ C correlations observed in triple resonance experiments (only 65% of the expected correlations in the [^{1}H,^{13}C,^{15}N]-TROSY HNCO and 45% in the [^{1}H,^{13}C,^{15}N]-TROSY HNCA experiments were observed) meant that only 20% of the amide cross peaks in triple resonance spectra gave the full complement of possible ^{13}C correlations (Plevin, 2003).

The possible reasons for this shortcoming include a common drawback of expressing proteins in D_2O media. In some cases, the usual strategy employed in order to back-exchange protons into labile positions (such as all amide groups), which consists of purifying the entirely deuterated protein in H_2O buffers, is not suitable for larger proteins. A significant number of cross peaks, *ca.* 35 in [^{2}H,^{15}N]-WT *Pa*DDAH spectra, were not recovered after the purification procedure, which lasted approximately four days. These amide groups presented exchange rates which were

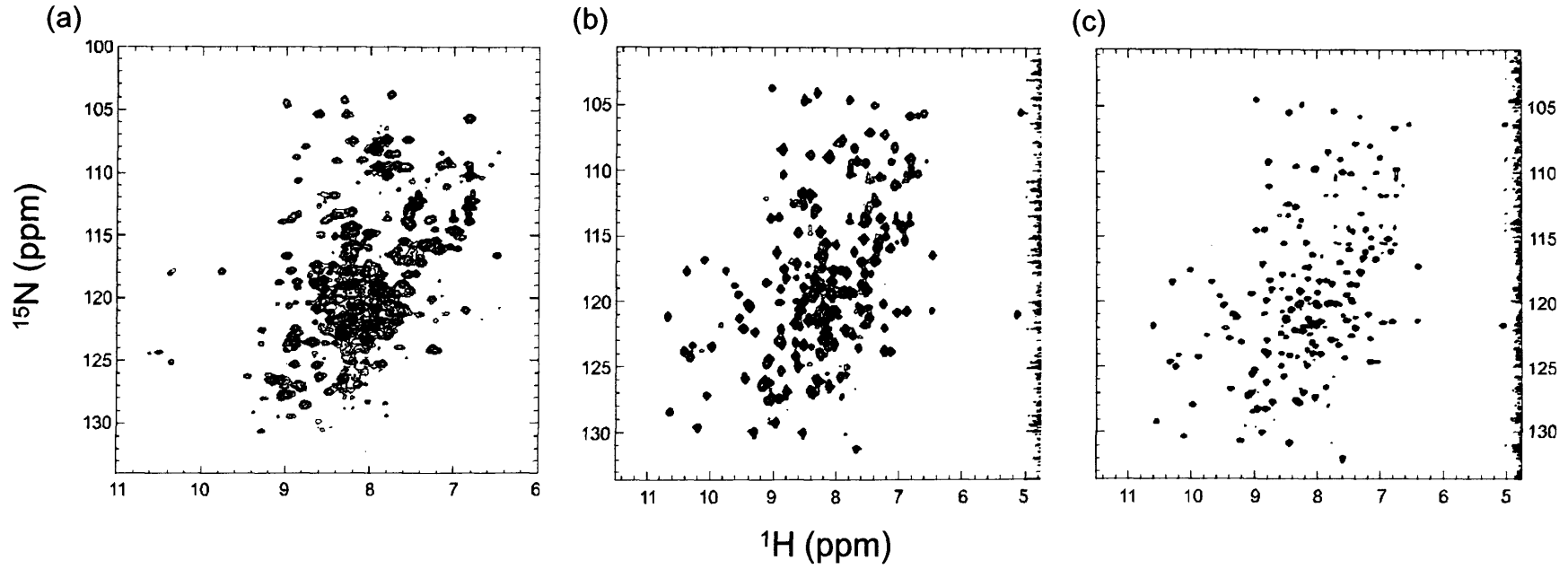

Figure 4.1. (a) [^{1}H, ^{15}N]-HSQC spectrum recorded of [^{1}H,^{15}N]-labelled WT *Pa*DDAH. Data matrices consisted of 512* (t_1) and 1024* (t_2) points (where n* refers to complex points) with acquisition times of 71.1 ms (t_1) and 102.4 ms (t_2). The experiment was recorded with 16 transients per t_1 increment. [^{1}H, ^{15}N]-HSQC (b) and [^{1}H, ^{15}N]-TROSY (c) spectra recorded of [^{2}H,^{15}N]-labelled WT *Pa*DDAH. Both [^{1}H, ^{15}N]-HSQC and [^{1}H, ^{15}N]-TROSY data matrices consisted of 400* (t_1) and 2048* (t_2) and points (where n* refers to complex points) with acquisition times of 100 ms (t_1) and 85 ms (t_2). A total of 32 transients were recorded per t_1 increment. All spectra were recorded at 25°C and at ^{1}H proton frequency of 500 MHz. Protein concentration was 0.8 mM and the affinity tag was still present. Adapted from Plevin, 2003.

extended to impracticable time scales and remained deuterated even after several weeks following purification (Plevin, 2003). These sites are evidently protected from solvent and implicated in stable secondary structure formation. Additional efforts to partially unfold and refold the enzyme to allow the complete back exchange to occur failed because of our inability to find the conditions necessary to improve the yield of active and refolded enzyme recovered. The refolding strategy, which proved successful for other proteins of the same size (Gardner *et al.*, 1998, Tugarinov *et al.*, 2002), was eventually abandoned as the yield was not satisfactory and not yet worth risking the loss of considerable amounts of [^{2}H,^{13}C,^{15}N]-labelled *Pa*DDAH protein.

Alternatively, the intrinsic monomer-dimer equilibrium of the protein could be at the origin of chemical and conformational exchange occurring at the dimer interface region, leading to substantial line broadening of the resonances. The dynamic nature of the interface packing with the concomitant line broadening of the resonances located in this region have been shown to occur in several other systems (Lukin *et al.* 2003, Yung *et al.*, 2003; Pfuhl *et al.*, 1999). In addition, *Pa*DDAH has also a flexible loop (L1) covering the active site region (Murray-Rust *et al.* 2001), for which residues may also experience exchange broadening of NMR signals.

If the backbone resonance assignments could be obtained for the monomeric *Pa*DDAH, a tentative identification of the missing $^{1}H^{N}$,^{15}N correlations in WT *Pa*DDAH [^{1}H,^{15}N]-TROSY spectra might offer a qualitative assessment of the regions of the protein undergoing conformational exchange and identify possible slowly solvent-exchanging amide groups. Such an analysis is discussed in detail in Section 4.4.

The study of monomeric PaDDAH by heteronuclear NMR

As outlined in the previous chapter, the more tractable engineered double mutant *Pa*DDAH provided an excellent platform for future NMR studies. The initial [^{1}H,^{15}N]-HSQC experiments recorded of [^{1}H,^{15}N]- and [^{2}H,^{15}N]-labelled R40E-R98H *Pa*DDAH (Figure 4.2) show a significant improvement of the spectral quality in terms of the increased number of visible cross peaks and of the uniformity of the S/N ratio when compared to the WT protein spectra. The central region of the spectrum, which is highly overlapped in the wild-type spectrum, also shows a clear improvement in signal resolution, which is further enhanced by deuteration (Figure 4.2b). Of particular

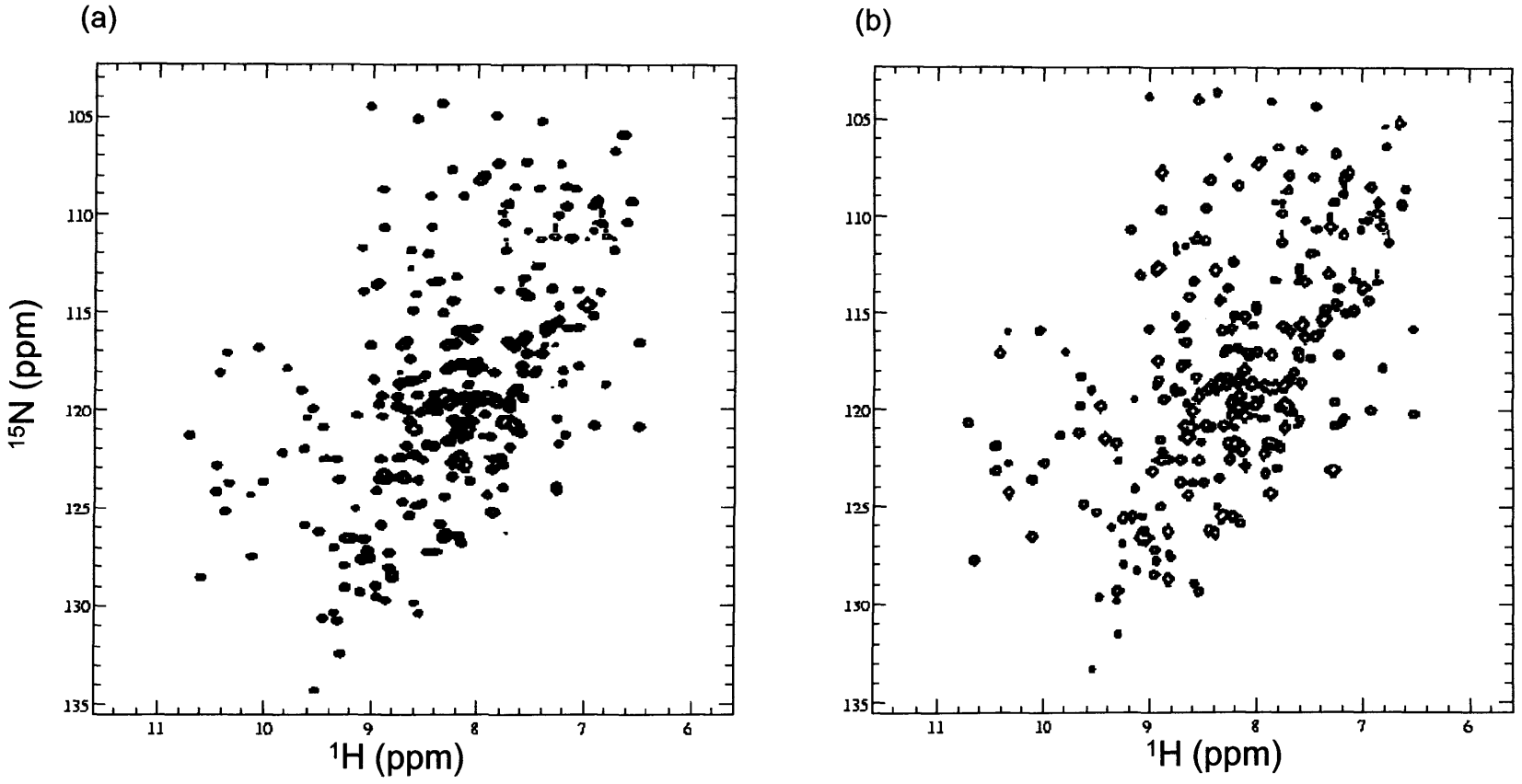

Figure 4.2. [^{1}H,^{15}N]-HSQC spectra recorded of (a) [^{1}H,^{15}N]-labelled R40E-R98H *Pa*DDAH, and (b) [^{2}H,^{15}N]-labelled R40E-R98H *Pa*DDAH. Data matrices consisted of 240* (t_1) and 1024* (t_2) points (where n* refers to complex points) with acquisition times of 75.0 ms (t_1) and 122.0 ms (t_2). The experiment was recorded with 8 transients per t_1 increment in (a) and 4 transients per t_1 increment in (b). Both spectra were recorded at 25°C and at ^{1}H proton frequency of 500 MHz. Protein concentration was 1 mM.

importance is the fact that the recombinant monomeric protein was less prone to incomplete reprotonation of exchangeable amide group protons than the WT protein when expressed in D_2O-containing media. The number of cross peaks present in the [^{1}H,^{15}N]-HSQC spectra of both protonated and perdeuterated samples is comparable (Figure 4.2). Unfortunately, a [^{1}H,^{15}N]-TROSY spectrum of the deuterated sample was not recorded which would have allowed an assessment of the overall improvement the combination of this experiment and deuteration might have provided. However, a [^{1}H,^{15}N]-TROSY spectrum of a protonated [^{1}H,^{15}N]-R40E-R98H *Pa*DDAH is shown in Figure 4.3b. From visual inspection alone, when compared to the HSQC spectra recorded at the same proton frequency (Figure 4.3a), the resolution is slightly increased in the central region of the spectrum. Despite this fact, the quality of the HSQC-based experiments is largely sufficient to discard the use of TROSY-based methods in the characterisation of this protein. Whereas for the dimeric WT protein, deuteration and the application of TROSY were essential to acquire high quality data, the smaller overall size of the monomer makes it tractable by more conventional NMR experiments.

4.3 ^{1}H, ^{13}C and ^{15}N backbone assignments of the monomeric *Pa*DDAH

Description of the triple resonance experiments

A number of heteronuclear 3D and 4D experiments have been proposed for the sequential assignment of uniformly isotopically labelled proteins (Sattler *et al.*, 1999). The pulse sequences generally rely on one- and two-bond heteronuclear scalar couplings (*J*) along the polypeptide chain which arise between spins. The establishment of sequential connectivity correlates the $^1H^N$ and ^{15}N nuclei of residue i with both its intra-residue (i) and inter-residue (i-1) $^{13}C_\alpha$ and side chain $^{13}C_\beta$ nuclei. The advantage of including the $^{13}C_\alpha$ and $^{13}C_\beta$ correlations is that the characteristic chemical shifts observed are often diagnostic of a specific amino acid. In this sense, information about the $^{13}C_\alpha$ and $^{13}C_\beta$ chemical shift values of the i-1, i and i+1 residues can potentially identify an amino acid triad within the known primary sequence of the protein (Grzesiek and Bax, 1993). Increasing the number of the connected spin systems increases the probability of correctly identifying a correspondence to a unique amino acid sequence within the protein.

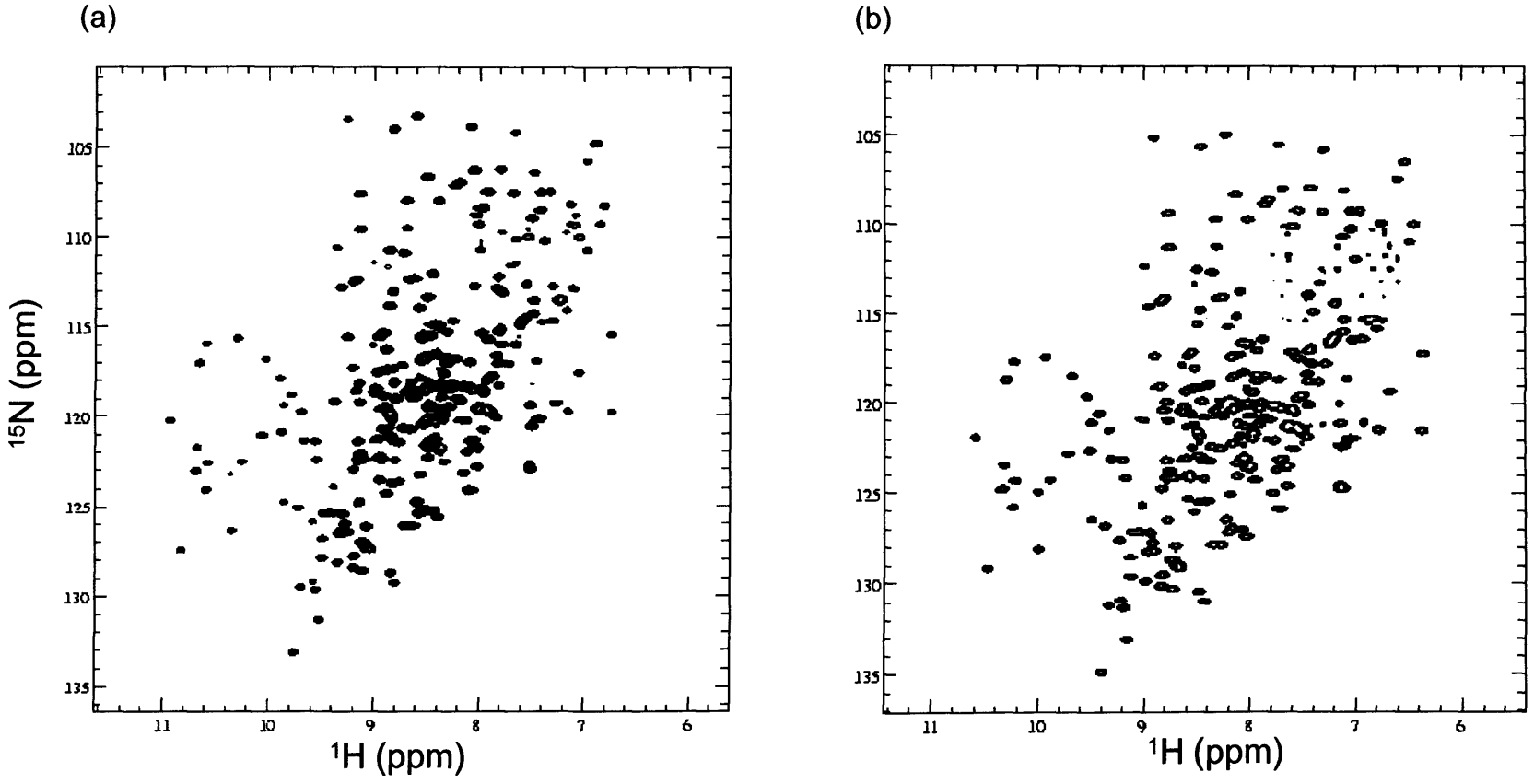

Figure 4.3. (a) [^{1}H,^{15}N]-HSQC and (b) [^{1}H,^{15}N]-TROSY spectra recorded of [^{1}H,^{15}N]-labelled R40E-R98H *Pa*DDAH. Data matrices consisted of 400*t_1 points in both (a) and (b) and 2048* t_2 points in (a) and 4096*t_2 in (b) (where n* refers to complex points) with acquisition times of 91.0 ms (t_1), 128.0 ms (t_2) in (a) or 204.8 ms (t_2) in (b). Both experiments were recorded with 32 transients per t_1 increment at 25°C and at ^{1}H proton frequency of 600 MHz. The protein concentration was 1 mM.

As mentioned before, the success of the triple resonance experiments for sequence-specific resonance assignments of proteins in the region of 30 kD relies on sufficiently long $^{13}C_{\alpha}$ T_2 and $^{1}H^{N}$ T_2 relaxation times to allow the establishment of the correlation with the $^{13}C_{\alpha/\beta}$ atoms. The most significant contribution to ^{13}C transverse relaxation comes from the strong dipolar coupling between the ^{13}C nucleus and the directly attached proton (Grzesiek *et al.*, 1993). Additionally, the dipolar coupling to the surrounding aliphatic protons has been shown to contribute significantly to $^{1}H^{N}$ T_2 relaxation. In large proteins these relaxation effects can lead to an absence of detectable C_{β} correlations in the HNCACB-based experiments (Wittekind and Mueller, 1993, Venters *et al.*, 1996, Garrett *et al.*, 1997). One way to circumvent the decline in sensitivity is to deuterate the protein and incorporate deuterium decoupling in the pulse schemes (Yamazaki *et al.*, 1994a,b).

In this context, the following 3D deuterium-decoupled gradient sensitivity-enhanced triple resonance experiments were selected for the assignment of monomeric *Pa*DDAH: HNCO, HNCA, HN(CO)CA, HN(CA)CB and HN(COCA)CB (Yamazaki *et al.*, 1994a,b; see Chapter 2, Section 2.5). The nomenclature of these experiments highlights the nuclei involved in the magnetisation transfer pathways (Figure 4.4). Spins whose chemical shifts are not evolved are denoted by parentheses. As shown in Figure 4.4 and summarised in Table 4.1, the magnetisation transfers in all these experiments are of the out-and-back type both originating and being detected on the H^{N} spin. More specifically, the combination of the HNCA, HN(CO)CA, HN(CA)CB and HN(COCA)CB experiments theoretically provide for each $^{1}H^{N}$,^{15}N resonance pair a set of four correlations HN $\rightarrow$ [$^{13}C_{\alpha}$(i), $^{13}C_{\alpha}$(i-1), $^{13}C_{\beta}$(i), $^{13}C_{\beta}$(i-1)] that then enables the sequential assignments of the backbone resonances.

Additional features of these pulse schemes, in the form adapted for studies on *Pa*DDAH, include the incorporation of pulsed field gradients in order to minimize artefacts and the residual water (Kay *et al.*, 1992a, Sattler *et al.*, 1999) and the use of an enhanced sensitivity approach to optimize S/N (Palmer *et al.*, 1991a). However, a detailed explanation of these features lies outside the scope of this thesis.

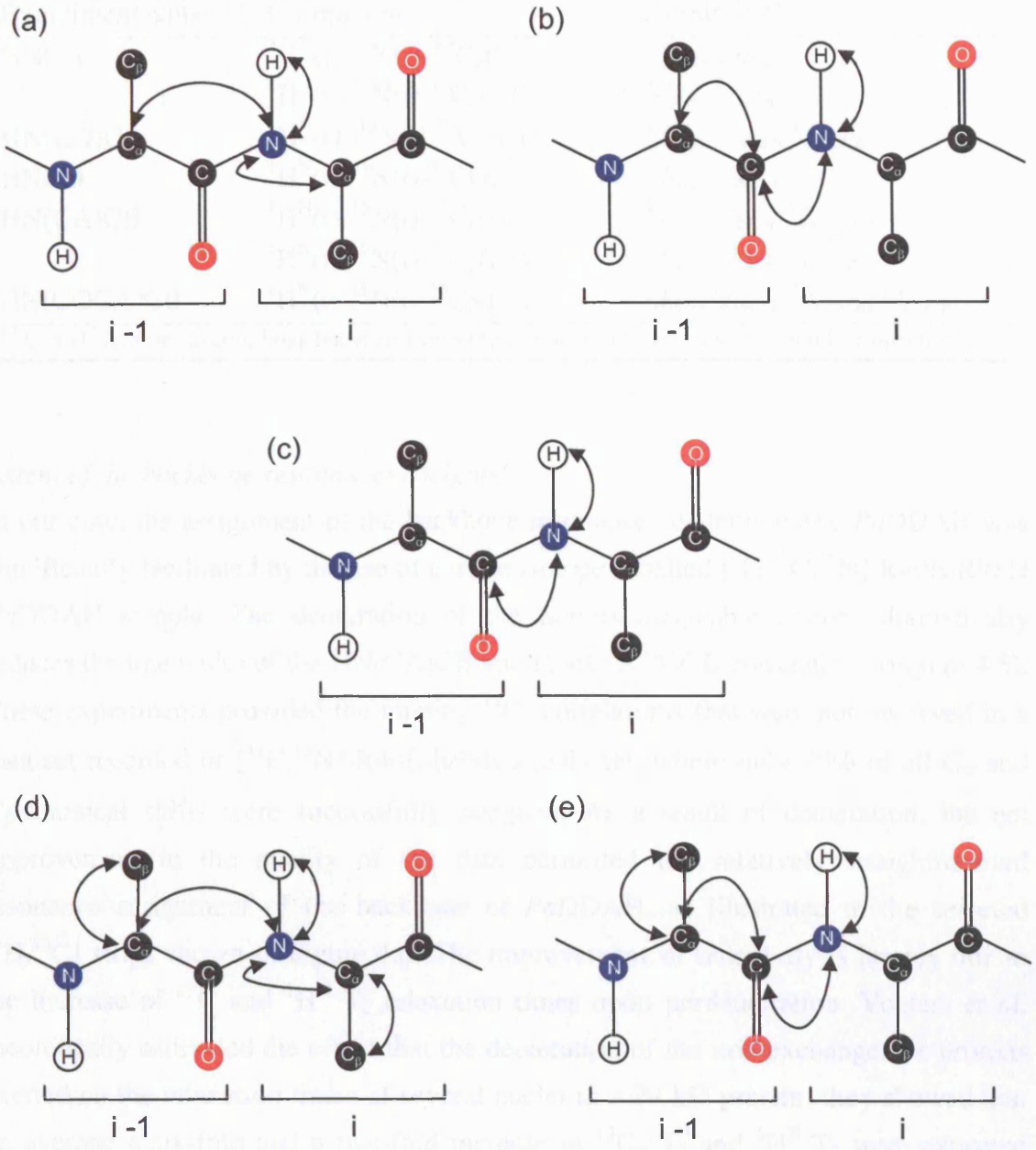

Figure 4.4. Experimental strategy for the resonance assignment of backbone ^{1}H, ^{15}N, $^{13}C_{\alpha}$, $^{13}C_{\beta}$ and CO nuclei in a dipeptide fragment showing residues i and i-1. Magnetisation transfer pathways in (a) HNCA, (b) HN(CO)CA, (c) HNCO, (d) HN(CA)CB and (e) HN(COCA)CB experiments. All five experiments shown are of the 'out-and-back' variety with magnetisation originating and ending on the amide proton. The direction of the magnetisation transfer is indicated by arrows. Spins whose chemical shift are not evolved are put in parentheses in the pulse sequence name.

Table 4.1. Summary of the spin correlations observed in the three-dimensional triple resonance experiments used for sequential assignment

Experiment name	Correlation	J coupling[a]
HNCA	$^{1}H^{N}(i)$-$^{15}N(i)$-$^{13}C_{\alpha}(i)$	$^{1}J_{NH}$, $^{1}J_{NC\alpha}$
	$^{1}H^{N}(i)$-$^{15}N(i)$-$^{13}C_{\alpha}(i-1)$	$^{1}J_{NH}$, $^{2}J_{NC\alpha}$
HN(CO)CA	$^{1}H^{N}(i)$ -$^{15}N(i)$-$^{13}C_{\alpha}(i-1)$	$^{1}J_{NH}$, $^{1}J_{NCO}$, $^{1}J_{COC\alpha}$
HNCO	$^{1}H^{N}(i)$-$^{15}N(i)$-$^{13}CO(i-1)$	$^{1}J_{NH}$, $^{1}J_{NCO}$
HN(CA)CB	$^{1}H^{N}(i)$-$^{15}N(i)$-$^{13}C_{\beta}(i)$	$^{1}J_{NH}$, $^{1}J_{NC\alpha}$, $^{1}J_{C\alpha C\beta}$
	$^{1}H^{N}(i)$-$^{15}N(i)$-$^{13}C_{\beta}(i-1)$	$^{1}J_{NH}$, $^{2}J_{NC\alpha}$, $^{1}J_{C\alpha C\beta}$
HN(COCA)CB	$^{1}H^{N}(i)$-$^{15}N(i)$-$^{13}C_{\beta}(i-1)$	$^{1}J_{NH}$, $^{1}J_{NCO}$, $^{1}J_{COC\alpha}$, $^{1}J_{C\alpha C\beta}$

[a] $^{1}J_{xy}$ and $^{2}J_{xy}$ refer to couplings between atoms (x,y) separated by one and two bonds, respectively.

Extent of the backbone resonances assigned

In our case, the assignment of the backbone resonances of monomeric *Pa*DDAH was significantly facilitated by the use of a triple isotope-labelled [^{2}H,^{13}C,^{15}N]-R40E-R98H *Pa*DDAH sample. The deuteration of the non-exchangeable amides dramatically reduces the line width of the HN(CA)CB and HN(COCA)CB correlations (Figure 4.5). These experiments provided the missing $^{13}C_{\beta}$ correlations that were not observed in a data set recorded of [^{13}C,^{15}N]-R40E-R98H *Pa*DDAH, where only 70% of all C_{α} and C_{β} chemical shifts were successfully assigned. As a result of deuteration, the net improvement in the quality of the data permitted the relatively straightforward resonance assignment of the backbone of *Pa*DDAH, as illustrated in the selected [^{1}H,^{13}C] strips shown in Figure 4.6. The improvement in sensitivity is largely due to the increase of ^{13}C and $^{1}H^{N}$ T_2 relaxation times upon perdeuteration. Venters *et al.* theoretically estimated the effect that the deuteration of the non-exchangeable protons exerted on the relaxation times of several nuclei in a 29 kD protein; they showed that on average a six-fold and a two-fold increase in $^{13}C_{\alpha}$ T_2 and $^{1}H^{N}$ T_2 were expected respectively (Venters *et al.*, 1996). These values have been confirmed experimentally for different proteins (Markus *et al.*, 1994, Garrett *et al.*, 1997, Gardner *et al.*, 1998).

The analysis of the triple resonance experiments recorded of [^{2}H,^{13}C,^{15}N]-R40E-R98H *Pa*DDAH allowed identification and the sequence-specific resonance assignments for 239 out of the 241 possible backbone ^{15}N and $^{1}H^{N}$ resonances (the *Pa*DDAH sequence has 254 residues of which 13 are prolines, however the construct has seven additional N-terminal residues remaining after cleavage of the hexa histidine-tag, of which five

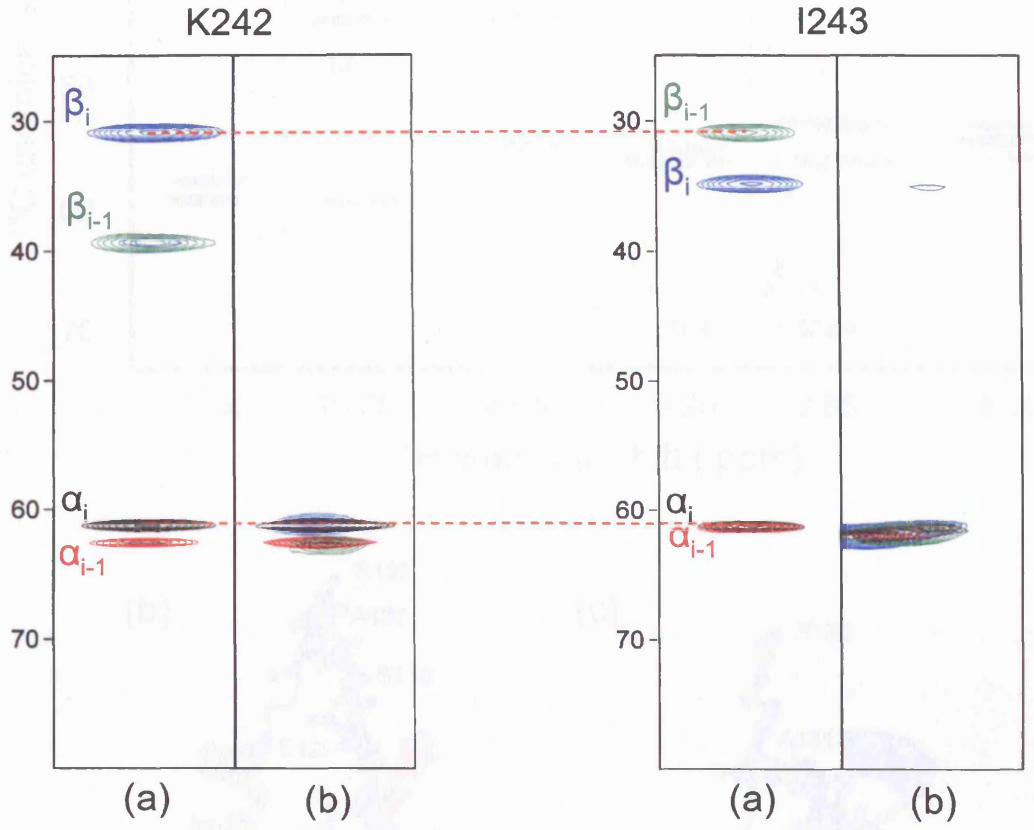

Figure 4.5. (a) [^{1}H,^{13}C] strip plots of the HNCA (black), HN(CO)CA (red), HN(CA)CB (blue) and HN(COCA)CB (green) spectra of [^{2}H,^{13}C,^{15}N]-labelled R40E-R98H *Pa*DDAH. (b) [^{1}H,^{13}C] strip plots of the HNCA (black), HN(CO)CA (red), HNCACB (blue) and CBCA(CO)NH (green) spectra of [^{13}C,^{15}N]-labelled R40E-R98H *Pa*DDAH. The strips show examples of $^{1}H^{N}$,^{15}N $\rightarrow$ ^{13}C correlations for residues K242 and I243. Cross peaks are labelled as originating from α or β, intra- i and inter-residue i-1 carbon atom correlations. Note: the HN(CA)CB and the HN(COCA)CB experiments were only acquired for the [^{2}H,^{13}C,^{15}N]-labelled protein, in contrast to the HNCACB and the CBCA(CO)NH experiments acquired for the [^{2}H,^{13}C,^{15}N]-labelled protein (not discussed in the text).

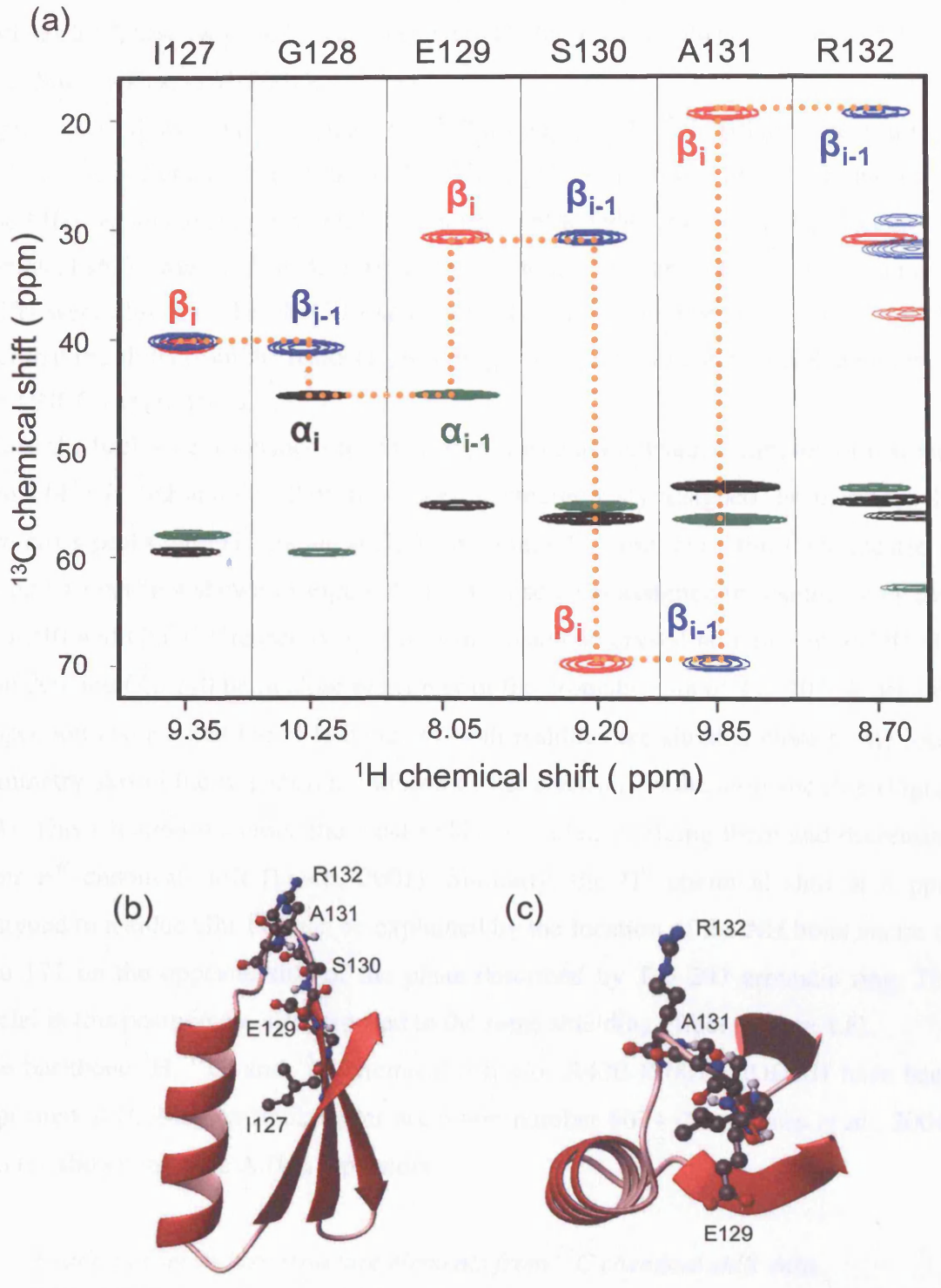

Figure 4.6. (a) Strip plots of the HNCA (black), HN(CO)CA (green), HN(CA)CB (red) and HN(COCA)CB (blue) spectra showing examples of $^1H^N$,$^{15}N \rightarrow {}^{13}C$ correlations from residues I127 to R132. Cross peaks are labelled as originating from α or β, intra- i and inter-residue i-1 carbon atom correlations. The dotted orange line depicts the sequential assignment of selected $^1H^N$,$^{15}N \rightarrow {}^{13}C_{\beta i}$ and $^{13}C_{\beta i-1}$ correlations. (b) Lateral and (c) plane view of the 3D structure of PaDDAH module 3. The residues for which $^1H^N$, $^{15}N \rightarrow {}^{13}C$ resonances are shown are depicted in a ball-and-stick model and follow CPK colouring.

have also been assigned). Definitive assignments have not been obtained for Ser 21 and His 22. These two residues are located in the large L1 loop which covers the active site (Murray-Rust *et al.* 2001).

Figure 4.7 shows an assigned 2D ^{13}C-decoupled ^{1}H-^{15}N HSQC spectrum of [^{2}H,^{15}N,^{13}C]-labelled R40E-R98H *Pa*DDAH at pH 7.0, recorded at a ^{1}H frequency of 500 MHz. From the assigned amide resonances, 99%, 99%, and 93% of C_α, C_β and CO chemical shifts were obtained, respectively. A number of the CO carbon-13 chemical shifts were absent in the HNCO experiment. This lower number is explained by the fact that the shifts from the residues preceding the 13 prolines cannot be obtained from the HNCO experiment alone.

All of the backbone resonances for the active site catalytic triad, composed of residues Glu 114, His 162 and Cys 249, have been unambiguously assigned. Furthermore the two cross peaks with H^N chemical shifts at around 4.8 ppm and a third one located at around 4 ppm (not shown in Figure 4.7) have also been assigned to residues Asn 209, Glu 210 and Glu 171 respectively. From the available crystal structure of *Pa*DDAH, Asn 209 and Glu 210 lie in close proximity of the aromatic ring of Trp 207. A detailed inspection reveals that the N-H bonds of both residues are situated close to the local symmetry axis of the ring current induced by the electrons in the aromatic ring (Figure 4.8). This situation increases the local fields of nuclei, shielding them and decreasing their H^N chemical shift (Levitt, 2001). Similarly, the H^N chemical shift at 4 ppm assigned to residue Glu 171 can be explained by the location of the NH bond vector of Glu 171 on the opposite side of the plane described by Trp 207 aromatic ring. The nuclei in this position are also exposed to the same shielding effect (Figure 4.8).

The backbone ^{1}H, ^{13}C, and ^{15}N chemical shifts of R40E-R98H *Pa*DDAH have been deposited at BioMagResBank under accession number 6074 (Magalhães *et al.*, 2004) and are shown in Table A.II in Appendix.

Identification of secondary structure elements from ^{13}C chemical shift data

It is well established that nuclei which are part of a secondary structure element show specific chemical shift changes in relation to their standardised random coil values (Wishart and Sykes, 1994). Namely, the chemical shifts of C_α nuclei experience an upfield shift when present in β-sheets and extended structures and a downfield shift

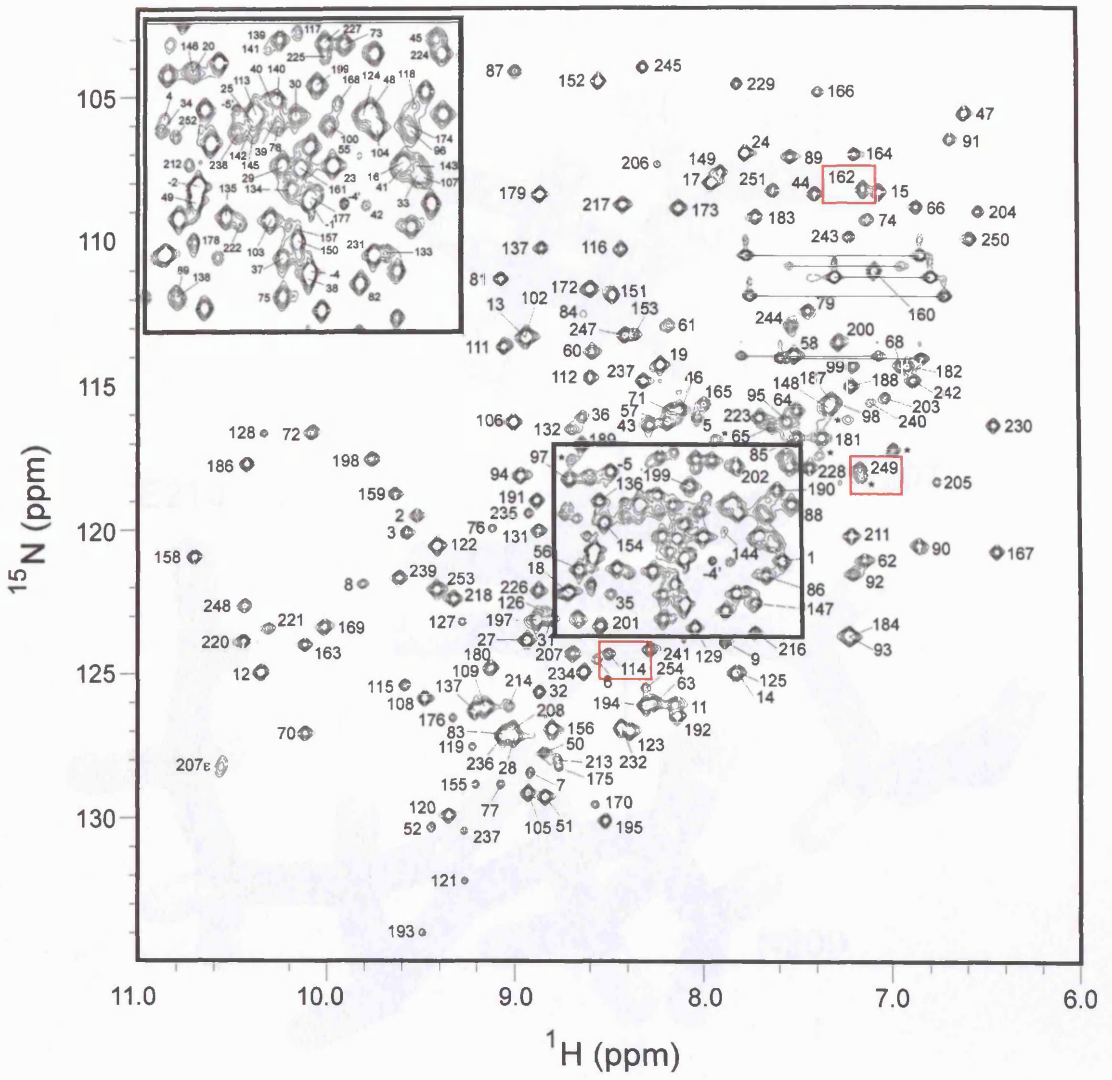

Figure 4.7. Assigned 2D ^{13}C-decoupled [^{1}H,^{15}N]-HSQC spectrum of [^{2}H,^{13}C,^{15}N]-labelled R40E-R98H *Pa*DDAH recorded on a 500 MHz Varian UNITYplus spectrometer at 298 K. Cross-peaks assigned by an asterix (*) are aliased arginine and lysine side-chain amide resonances. Cross peaks highlighted in a red box correspond to residues of the catalytic triad.

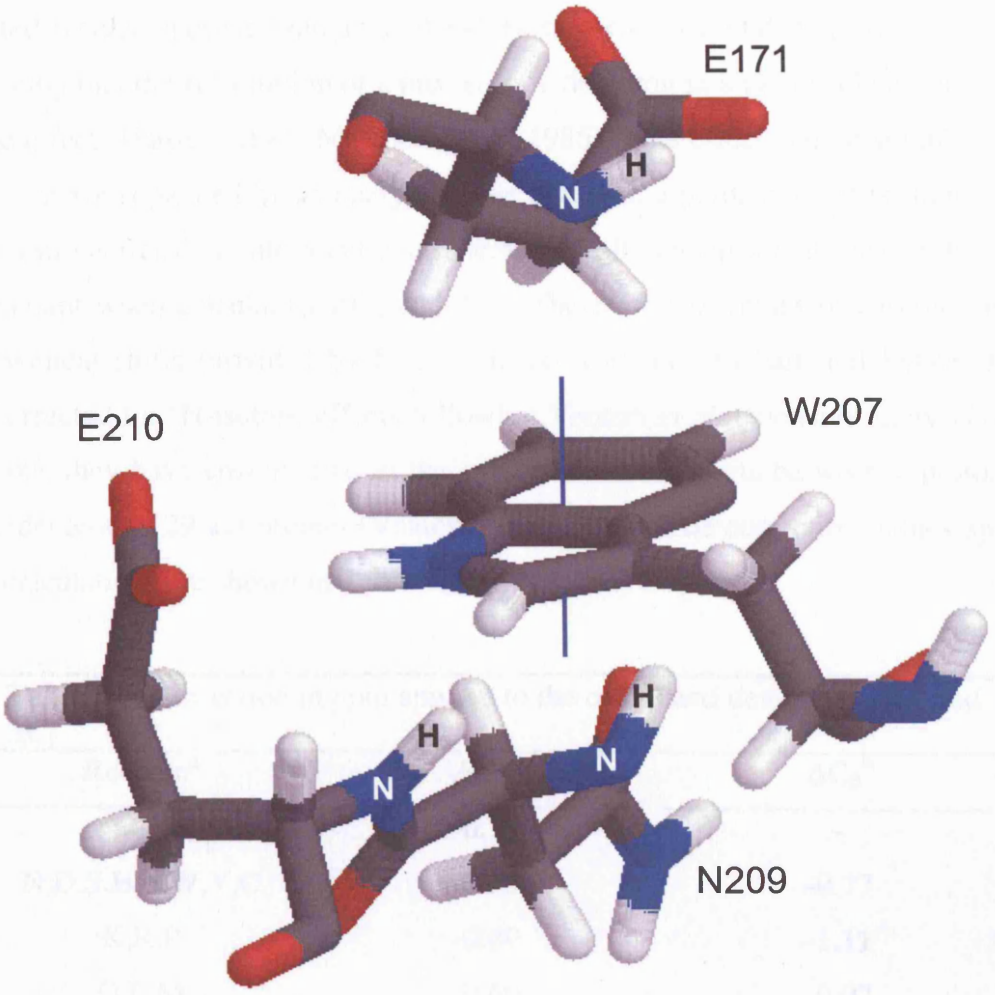

Figure 4.8. 3D wireframe diagram of residues Glu 171, Trp 207, Asn 209 and Glu 210 of *Pa*DDAH (pdb code: 1h70). The N-H bond vectors of E171, N209 and E210 lie close to the local symmetry axis described by Trp 207 aromatic ring (blue line). Atoms follow CPK colouring.

when present in α-helices. Conversely, C_β nuclei chemical shifts display the opposite correlation with local secondary structure (Spera and Bax, 1991). In this respect, in order to estimate the chemical shift deviations which in groups are diagnostic of the presence of secondary structure elements, the experimental $^{13}C_\alpha$ and $^{13}C_\beta$ chemical shifts obtained from the *Pa*DDAH assignment data should be subtracted from the tabulated residue-specific random coil values to yield ΔC_α and ΔC_β. However, it is noteworthy that the substitution of a proton for a deuteron in a carbon chain induces an isotope effect (Hansen, 1988, Majerski *et al.*, 1985). This effect can be significant for nuclei that are separated by as many as three bonds in a perdeuterated protein. These considerations translate into additional chemical shifts changes that have to be taken into account when calculating ΔC_α and ΔC_β. Therefore the values of carbon random coil chemical shifts provided by Sykes and co-workers (Wishart and Sykes, 1994) were corrected for ^{2}H-isotope effects following Venters *et al.* extensive analysis of the deviations they have encountered in the carbon chemical shifts between a protonated and perdeuterated 29 kD protein (Venters *et al.*, 1996). The correction values applied in our calculations are shown in Table 4.2.

Table 4.2. Correction in ppm applied to the calculated deviations ΔC_α and ΔC_β

Residue[a]	ΔC_α[b]	ΔC_β[b]
G	-0.39	-
N,D,S,H,F,W,Y,C	-0.55	-0.77
K,R,P	-0.69	-1.11
Q,E,M	-0.69	-0.97
A	-0.68	-1.00
I	-0.77	-1.28
L	-0.62	-1.26
T	-0.63	-0.81
V	-0.84	-1.20

[a]A one letter code is used to describe the residue in question. [b]$\Delta C_{\alpha/\beta} = {}^{13}C^{exp} - {}^{13}C^{rc}$, where exp and rc stand for the experimental $^{13}C_{\alpha/\beta}$ chemical shifts and the residue-specific random coil chemical shifts, respectively.

Since the deviations ΔC_α and ΔC_β are of similar magnitude and opposite signs whether in an α-helix or β-sheet, the difference between the two deviations (ΔC_α -ΔC_β)

enhances the correlation between the definition of the secondary structure and the chemical shift deviations. Where the carbon shifts were available, $\Delta C_\alpha - \Delta C_\beta$ values were then calculated for each *Pa*DDAH residue and smoothed according to:
$(\Delta C_\alpha - \Delta C_\beta)_{smoothed,i} = (\Delta C_{\alpha,i+1} + \Delta C_{\alpha,i} + \Delta C_{\alpha,i-1} - \Delta C_{\beta,i+1} - \Delta C_{\beta,i} - \Delta C_{\beta,i-1})/3$, where i is the residue number and $(\Delta C_\alpha - \Delta C_\beta)_{smoothed,i}$ represents an average chemical shift over three consecutive residues (Salzmann *et al.*, 2000). Figure 4.9 shows a plot of $(\Delta C_\alpha - \Delta C_\beta)_{smoothed}$ as a function of residue number. Values for which $\Delta C_\alpha - \Delta C_\beta > 1.4$ ppm identify likely helical regions and values for which $\Delta C_\alpha - \Delta C_\beta < -1.4$ ppm identify likely β-strand regions. The locations of the secondary structure elements predicted from the ^{13}C chemical shift data are, in general, well correlated with the α-helical and β-strand regions in the crystal structure of WT *Pa*DDAH. All five α-helices that form the five-fold ββαβ structural motifs are well delimited and share characteristic positive values of $(\Delta C_\alpha - \Delta C_\beta)_{smoothed}$. Nonetheless helix α4 which encompasses residues 184 to 188 remains ill defined by the NMR data. Inspection of the crystal structure however provides a good rationale for the low chemical shift values obtained, as this helix is rather small composed of only 3 residues (Figure 4.9b). Another point worth mentioning pertains to the analysis of the chemical shifts for helix α2. The NMR data suggest that helix α2 is in fact defined by two consecutive smaller helices. The ^{13}C chemical shifts from residues Ile 91 and Ile 92 deviate from the ideal values of α-helical behaviour and are more proximate to random coil values. This disagreement is in reality verified in the crystal structure which clearly shows that helix α2 is slightly tilted disturbing the hydrogen-bond network (Figure 4.9c). Additionally, the header of the pdb entry for the X-ray structure of *Pa*DDAH (pdb code: 1h70) defines helix α2 as two separate but consecutive smaller helices from residues 82 to 86 and 87 to 100 which further supports the notion that helix α2 exhibits a slight turn in line with the prediction made from the $(\Delta C_\alpha - \Delta C_\beta)_{smoothed}$ values.
Several other regions of the enzyme have been highlighted in Figure 4.9 as exhibiting positive $(\Delta C_\alpha - \Delta C_\beta)_{smoothed}$ values. These regions are two to three residues long and they have been outlined in the crystal structure as 3_{10}-helix turns and hence display chemical shift deviations higher than 1.4 ppm.
In turn, the residues located in β-strands are also well represented by the chemical shift data. They invariably exhibit the most negative values of $(\Delta C_\alpha - \Delta C_\beta)_{smoothed}$ which are indicative of hydrogen bonding. The only apparent discrepancy is that the data for residues 111-113 and 157-159 predict a β-sheet-like conformation for these regions,

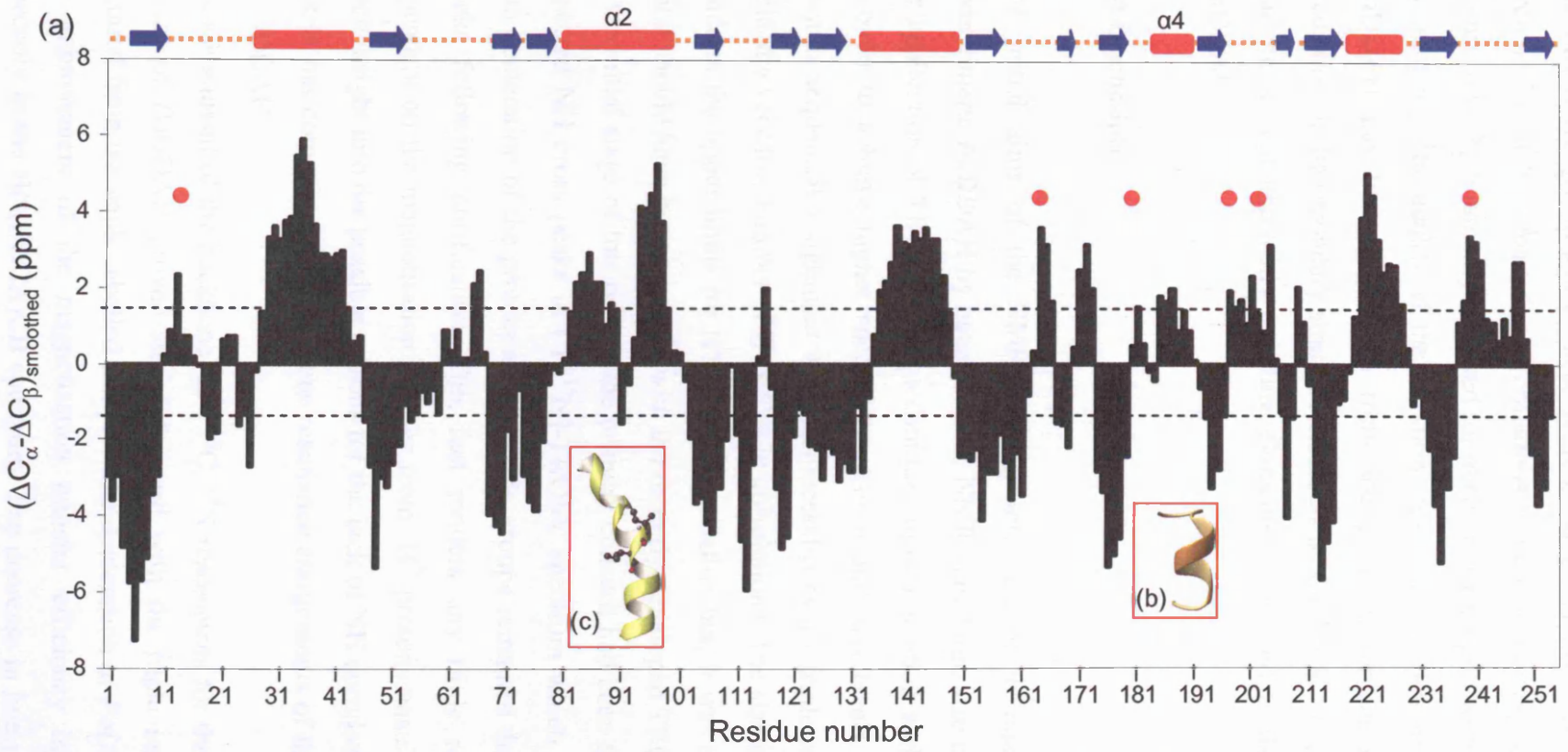

Figure 4.9. Identification of secondary structure elements of the monomeric *Pa*DDAH from ^{13}C chemical shift analysis. (a) $(\Delta C_\alpha - \Delta C_\beta)_{smoothed}$ values are plotted as a function of residue number. The dashed lines indicate the 1.4 ppm threshold value for the identification of secondary structure elements. Values higher or lower than 1.4 ppm or -1.4 ppm indicate the presence of α-helical or β-strand regions respectively. The locations of the α-helices and β-sheets in the X-ray structure are indicated as red rectangles and blue arrows at the top of the figure. The red dots indicate regions of the *Pa*DDAH structure which correspond to 3_{10}-helix turns. (b) Detail showing the 3D structure of the small helix α4 (pdb: 1h70) (c) Detail showing the tilt of helix α2. Residues Ile 91 and Ile 92 are shown in ball-and-stick and CPK colouring.

although they are not defined as such in the pdb file. However from the inspection of the *Pa*DDAH 3D structure, these regions adopt an extended conformation. Nuclei present in such extended conformations generally display the characteristic ^{13}C chemical shifts commonly observed for nuclei present in β-sheets.

In summary, the quality of the backbone resonance assignments for the monomeric *Pa*DDAH has been confirmed from the good agreement reached between the prediction of the secondary structure elements by the NMR ^{13}C chemical shift analysis and the regular secondary structure elements recognised in the crystal structure of *Pa*DDAH.

4.4 Discussion

The initial aims of the NMR investigations sought to characterise the 58 kD homodimeric *Pa*DDAH by heteronuclear NMR. The large size of the protein required the application of TROSY-based techniques in combination with uniform deuteration in order to achieve higher spectral resolution and S/N levels. However the efforts towards sequential assignment were unsuccessful as the triple resonance experiments yielded a very low number of H-N-C spin correlations. The size of WT *Pa*DDAH does reside in the upper limits for NMR characterisation but, it was reasonable to suppose that it should have been tractable with the recently developed TROSY pulse sequences. In the initial stage of this project the primary concern had been the lack of a subset of expected NH cross peaks in [^{1}H,^{15}N]-TROSY spectrum which was only aggravated with deuteration of the protein as the amide groups remained deuterated after several weeks following purification. This fact renders any triple resonance experiment dependent on the magnetisation transfer from $^{1}H^{N}$ protons unsuitable. In this respect, much insight into the possible reasons for the lack of NH correlations for the wild-type protein has come from the backbone resonance assignments of the monomeric variant of *Pa*DDAH.

The assignment of the backbone $^{1}H^{N}$, ^{13}C, ^{15}N resonances for the R40E-R98H double mutant of *Pa*DDAH proved straightforward with the triple resonance experiments acquired from the triple labelled sample. The deuteration of *Pa*DDAH was crucial for the improvement of the magnetisation transfer efficiency in these experiments, especially in the HN(COCA)CB spectrum. The decrease in line width was substantial and allowed us to assign 99% of the $^{1}H^{N}$, ^{13}C, ^{15}N resonances as only two residues

remained unassigned (Ser 21 and His 22). The smaller overall size of the monomer has dramatically facilitated the study of the protein, as the common drawback of deuterating the protein was single-handedly solved with no adjustments to the purification protocol. Furthermore, the prediction of the locations of the secondary structure elements from the analysis of the ^{13}C chemical shift data are well correlated with the positions of the respective structural elements in the crystal structure of the *Pa*DDAH homodimer (Murray-Rust *et al.*, 2001), further supporting the reliability of the data that enabled us to obtain manually the unambiguous assignments of the backbone resonances for R40E-R98H *Pa*DDAH.

The backbone resonance assignments of the monomeric *Pa*DDAH enabled the attempt of a "conservative" assignment of the $^{1}H^{N}$,^{15}N resonances of the WT *Pa*DDAH. In order to assess which amide NH groups were likely to remain deuterated upon WT *Pa*DDAH expression in D_2O media, two [^{1}H,^{15}N]-TROSY spectra of [^{1}H,^{15}N]- and [^{2}H,^{15}N]-WT *Pa*DDAH were superimposed and their resonances compared (Figure 4.10a). The cross peaks that were absent in the 2D spectra of the perdeuterated protein were tentatively identified by reference to the assignments of the monomer (Figure 4.10b). Only residues that were confidently identified with this procedure were plotted on the crystal structure of the homodimer (Figure 4.10c). Interestingly, the regions of the protein that remained protected from solvent exchange are confined to the set of three β-strands of modules 2 through 5 and to several residues in helix-α3 of module 3. All of these protected residues are located in regular secondary structure elements, in particular in the β-sheets which are more likely to be involved in the formation of a stable hydrogen bond network and hence lead to slow reprotonation of the amide positions. Intriguingly, the amide groups responsible for the hydrogen bonds in the β-sheet of module 1 were not identified as being protected from solvent. This finding is counterintuitive considering that the β-strands 1A, 1B and 1C form the interfacial contact of the homodimer and the extended antiparallel β-sheet across protomers. A potential explanation comes from the mapping of the missing resonances in [^{1}H,^{15}N]-TROSY spectra of the protonated WT *Pa*DDAH. In this case, the [^{1}H,^{15}N]-TROSY spectra of ^{15}N-labelled WT and R40E-R98H-*Pa*DDAH were superimposed (Figure 4.11). In a similar manner, residues for which resonances were present in the 2D spectrum of the monomeric enzyme but absent in the wild-type spectrum were plotted on the crystal structure (Figure 4.11c). By this assessment 34 $^{1}H^{N}$, ^{15}N missing cross peaks were confidently identified as missing. The absent correlations for the WT

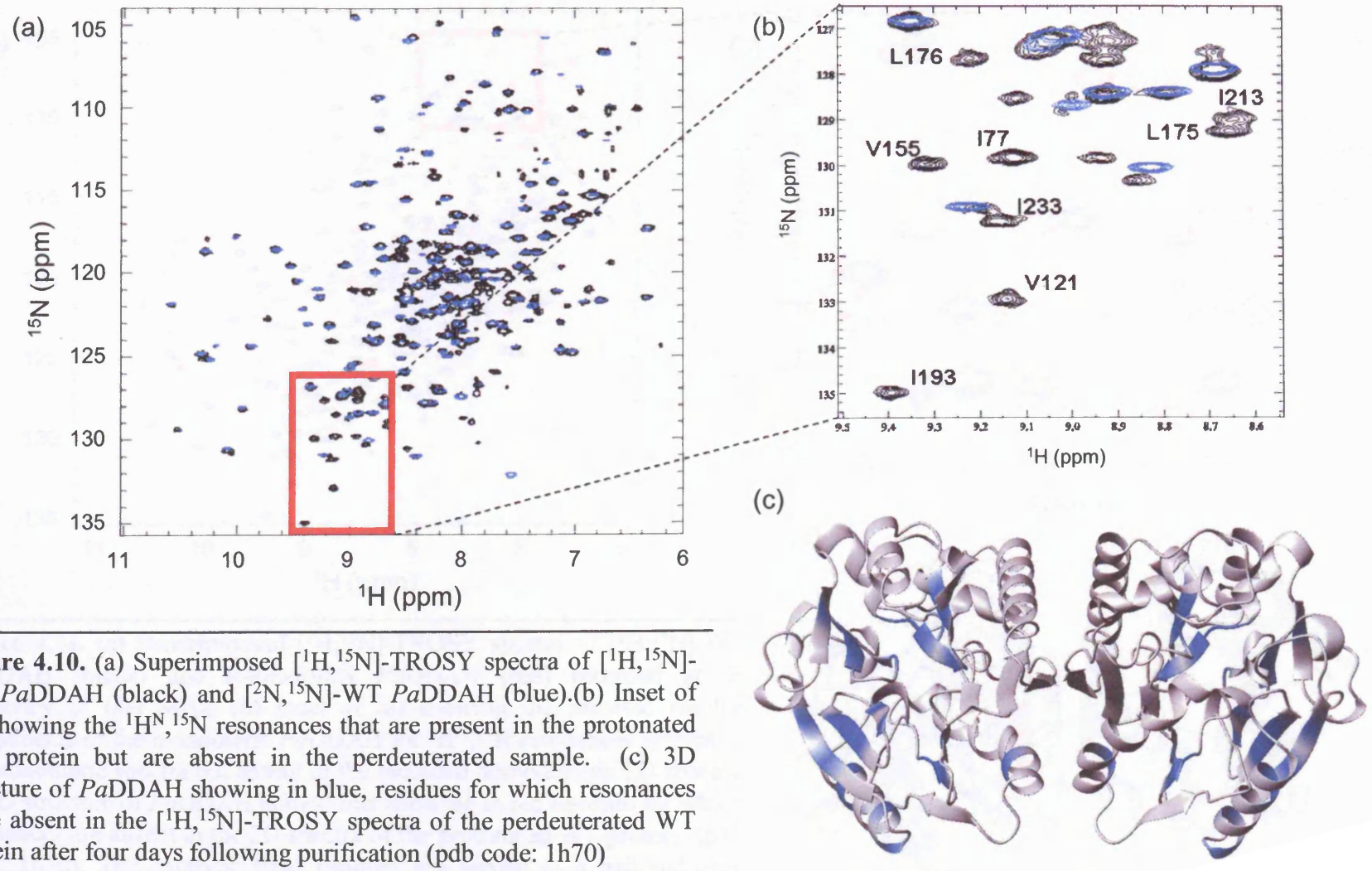

Figure 4.10. (a) Superimposed [^{1}H,^{15}N]-TROSY spectra of [^{1}H,^{15}N]-WT *Pa*DDAH (black) and [^{2}N,^{15}N]-WT *Pa*DDAH (blue).(b) Inset of (a) showing the $^{1}H^{N}$,^{15}N resonances that are present in the protonated WT protein but are absent in the perdeuterated sample. (c) 3D structure of *Pa*DDAH showing in blue, residues for which resonances were absent in the [^{1}H,^{15}N]-TROSY spectra of the perdeuterated WT protein after four days following purification (pdb code: 1h70)

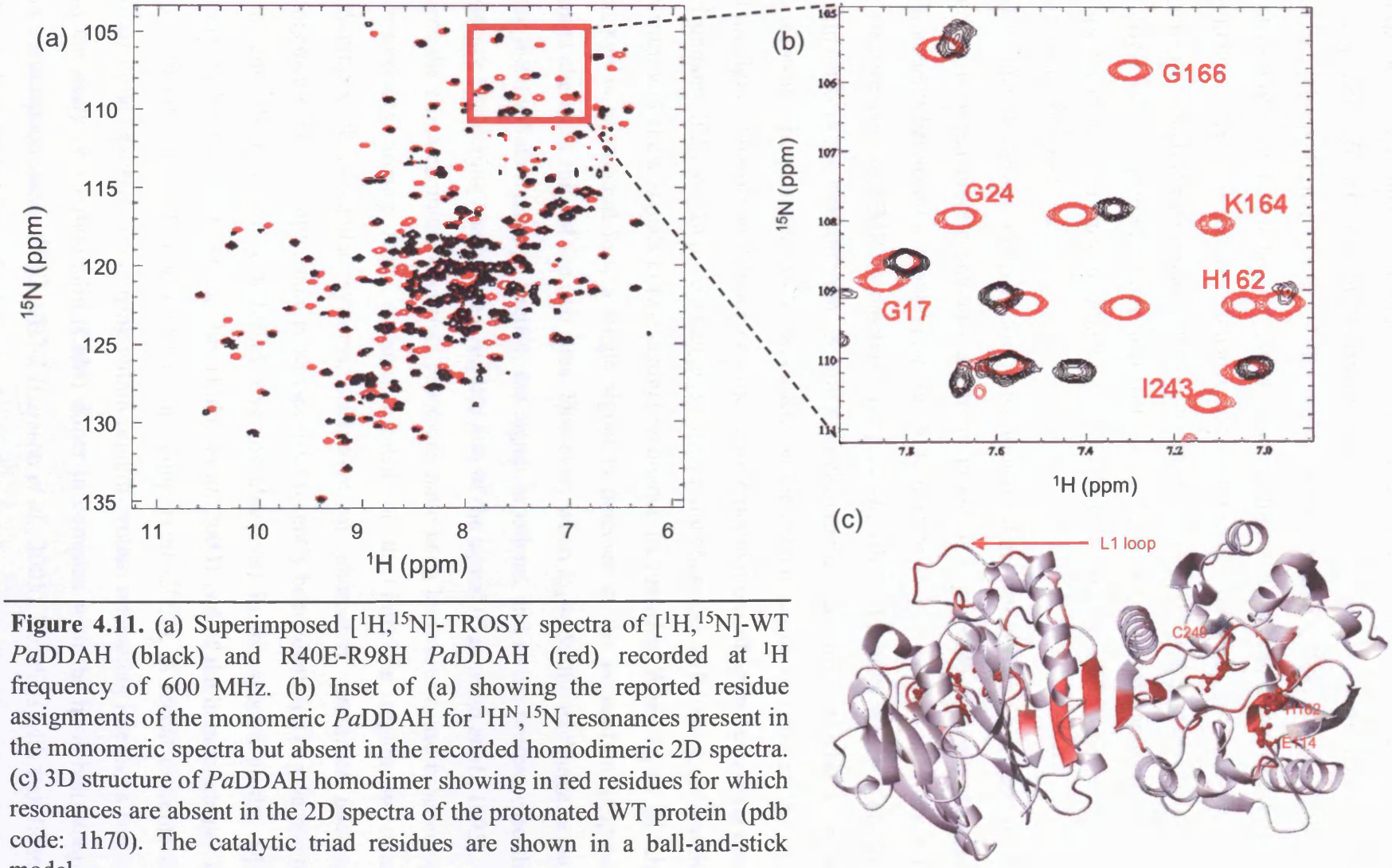

Figure 4.11. (a) Superimposed [^{1}H,^{15}N]-TROSY spectra of [^{1}H,^{15}N]-WT *Pa*DDAH (black) and R40E-R98H *Pa*DDAH (red) recorded at ^{1}H frequency of 600 MHz. (b) Inset of (a) showing the reported residue assignments of the monomeric *Pa*DDAH for $^{1}H^{N}$,^{15}N resonances present in the monomeric spectra but absent in the recorded homodimeric 2D spectra. (c) 3D structure of *Pa*DDAH homodimer showing in red residues for which resonances are absent in the 2D spectra of the protonated WT protein (pdb code: 1h70). The catalytic triad residues are shown in a ball-and-stick model.

protein correspond to residues which are located in module 1 and part of the extended β-sheet across subunits. More importantly, in addition to the interface region, none of the cross peaks for residues of the active site catalytic triad or the residues in the L1 loop are present in the spectra of the protonated homodimer. This examination has confirmed our initial surmise that chemical and conformational exchange occurs at the interface in the intermediate exchange rate regime. Coupled to this behaviour, it appears as though the monomer-dimer equilibrium of the wild-type protein could be at the origin of a concerted motion that influences the exchange broadening for amide groups located in the surrounding loops including the region where the active site residues lie (Figure 4.11c).

Due to the dynamics implicit for a monomer-dimer equilibrium, residues in an interface between protein subunits are more prone to experience different chemical environments between the monomeric and dimeric states, which is often detrimental to the measurement of NMR parameters (Tong *et al.*, 2005). In cases where a nucleus exchanges between two different chemical environments, the difference between the two resonance frequencies, $\Delta\upsilon$, is dependent on the exchange rate k_{ex} between the sites and the signal intensity will depend on the relative populations of spins occupying each environment. If $k_{ex} << \Delta\upsilon$ the exchange is slow on the chemical shift time scale and the spectrum will show signals at two distinct resonance frequencies. When $k_{ex} >> \Delta\upsilon$, the exchange is considered fast, a single signal is detected at the population weighted average chemical shift of the two sites. However, when $k_{ex} \approx \Delta\upsilon$ the exchange occurs in the intermediate exchange regime, the signal broadens, and in the present case the resonance broadening results in a complete loss of the signal (Canavagh *et al.*, 1995).

Intermediate rate chemical exchange processes have been known to cause broadening of resonances arising from residues located at the interface region in other multicomponent molecular systems, indicative of alternative protomer packing arrangements. For example, such processes have recently been identified to give rise to significant $^{15}N,^{1}H$ cross peak broadening (or absences) in the spectra of the $\alpha_2\beta_2$ tetramer of human adult hemoglobin (Lukin *et al.* 2003) and of the homopentamer B subunit (VTB) of the *E. coli* O157 toxin (Yung *et al.* 2003). A recent example of conformational exchange occurring within protein-protein interacting interfaces comes from the study of a calmodulin (CaM) dimer in complex with the basic-helix-loop-helix transcription factor SEF2-1/E2-2 (Larsson *et al.*, 2005). Residues with significant exchange broadening contributions are located not only within the dimer interface of

the calmodulin subunits but also within the binding interface of the CaM:SEF2-1 complex. However in all the systems cited above the exchange rates between the alternative states could be quantified.

Our results for the WT *Pa*DDAH only reconcile instances where the extent of the exchange broadening is so extreme that it hampers further NMR studies. In this sense, the monomerisation of the enzyme has been central to effectively abolish the deleterious consequences of conformational exchange at the interface region and surrounding loops. The additional "recovered" cross peaks in the [^{1}H,^{15}N]-HSQC spectra of the double mutant allowed for the successful assignment of the backbone resonances of the monomeric *Pa*DDAH and helped us to better understand the limitations of the NMR data for the WT protein. Altogether, the monomerisation of *Pa*DDAH has provided a protein variant that is an excellent platform for further studies of the properties of the enzyme.

The following chapter will focus on the characterisation of the dynamics of the apo-form of the monomeric *Pa*DDAH enzyme. ^{15}N relaxation studies of R40E-R98H *Pa*DDAH have been performed at three different magnetic field strengths and the data have been analysed with the model-free formalism to enable the quantification of the internal motion parameters. Recently, much insight into enzyme function has come from the linkage between the motions of the protein backbone in the microsecond to millisecond timescales. In this vein, the implementation of novel ^{15}N relaxation dispersion experiments has also been attempted to better assess the internal motion of *Pa*DDAH.

Chapter 5

^{15}N relaxation studies of the monomeric *Pa*DDAH

Abstract

The study of protein internal dynamics over the years has established an intimate association between protein motion and biological function. Central to this effort has been the application of NMR methodology as a powerful technique to characterise motional processes at atomic resolution. In particular, the measurement of nuclear spin relaxation rates provides information on protein movement over a wide range of timescales. As part of the efforts to investigate the intrinsic dynamic properties of the monomeric *Pa*DDAH and the possible functional role of these motions in enzyme catalysis, the quantification of the internal motions of the apo- state of *Pa*DDAH were sought through established ^{15}N NMR relaxation experiments. A vast array of protocols have been developed for the analysis of these parameters and the data have been interpreted through the "model-free" formalism of Lipari and Szabo (Lipari and Szabo, 1982a,b) and through reduced spectral density mapping (Farrow *et al.*, 1995). A general view that arises from the interpretation of these results is the overall stability of the β/α propeller fold as evidenced by the relatively high average value of the generalised S^2 order parameter of the NH bond vector (~0.87). However, local motions in the fast pico- to nano-second timescales have been observed in the L1, L3, L6 and L7 loops. All these loops form the entrance of the active site cavity. With regards to motions in the micro- to milli-second timescale, they are mainly located in the active site region and surrounding residues that contact the ligand as observed in the *Pa*DDAH crystal structure in complex. Independent confirmation of the results obtained from model-free analysis was sought through a linear regression protocol and reduced spectral density mapping. Finally, an analysis of the biological significance of these motions is presented in terms of a comparative study with other enzymes.

5.1 Coupling NMR spin relaxation and protein dynamics

NMR spin relaxation

The study of protein dynamics by NMR is inherently linked to the relaxation of spin systems. In simple terms, spin relaxation is the process through which a given excited state of spin magnetisation returns to the equilibrium position through the interaction with the thermal molecular environment. The analysis of the rates at which spin relaxation occurs can effectively report on the internal dynamics of proteins occurring at a wide range of timescales. Since the majority of experiments used to probe the dynamics of a protein backbone benefit from the relaxation rates of the ^{15}N spin, the treatment of nuclear relaxation theory presented here will only be focused upon the headline concepts of the relaxation properties of this nucleus.

When considering an excited two-spin system (^{1}H-^{15}N) in the presence of an applied B_0 field, several relaxation mechanisms exist to bring the system back to equilibrium with its surroundings. Generally, these mechanisms include the longitudinal relaxation defined by the rate constant R_1, the transverse relaxation defined by the rate constant R_2 and the steady state $\{^{1}H\}$-^{15}N heteronuclear Overhauser effect (NOE).

The longitudinal relaxation is the mechanism through which H or N spin magnetisation realign along the direction of the B_0 field, by convention, the z-axis of the system. In this sense, ^{15}N R_1 measures the rate at which the excited N magnetisation is returned to the equilibrium position, re-establishing net N_z magnetisation. This process is dictated by the Boltzman distribution, which states that at equilibrium, the lowest energy level will have the highest probability of being occupied. In this case, the energy state with the lowest energy is the one in which the magnetic moments of the ^{15}N nuclei are aligned along the z-axis accounting for the net N_z magnetisation (for more details, see Keeler, 2005). In turn, transverse relaxation describes the process in which the magnetisation H or N present in the x-y transverse plane decays to the equilibrium position, ultimately driving H_z or N_z to zero. Hence, the ^{15}N transverse relaxation rate R_2 measures the time dependent loss of $N_{x,y}$ magnetisation in the transverse plane. The third relaxation parameter commonly measured, the steady state $\{^{1}H\}$-^{15}N NOE, is dependent upon an additional mechanism of relaxation termed cross-relaxation. This mechanism gives rise to z-magnetisation transfer from one spin to the other at a specific transfer rate defined as σ_{NH}. Notably, the z-magnetisation of the ^{15}N spin will be affected by the z-magnetisation of the ^{1}H spin. This transfer results in what is

classically referred to as an enhancement of the signal recorded from the heteronuclei. However in the actual case of the ^{15}N nucleus, the magnetisation transfer results in a suppression of the NMR signal which is ultimately the NOE value that is quantified.

Having defined the three relaxation parameters that are commonly measured for the ^{15}N nucleus, it is important to understand what causes the spins to relax. In this sense, the major source of spin relaxation arises from the presence of random localised fields in the surroundings that can act as pulses, affecting the spins of the sample locally, eventually bringing the magnetisation to equilibrium. These local fields are commonly generated from the dipolar interaction between the magnetic moments of two different nuclear spins. The strength of this interaction is dependent upon the gyromagnetic ratios γ of the spin generating the local field and the spin experiencing it. The larger γ is, the larger the interaction will be between the spins. In addition, the strength is also dependent upon the distance r between the spins (the interaction decays as a function of $1/r^3$) and upon the orientation of the vector between the dipoles with respect to the applied magnetic field. Alternatively, a second source of random localised fields is the chemical shift anisotropy (CSA). In this case, the electrons in motion around the nucleus are responsible for generating the local fields. For most molecules, the size of these electron-induced local fields is dependent upon the orientation of the molecule relative to the applied external field, which gives this interaction its anisotropic character. The resulting field felt at the nucleus is then the sum of the applied external field and of the electron-induced local fields. This is at the origin of the chemical shift seen in NMR spectra.

In order for the dipolar interaction and the CSA to induce relaxation, the local magnetic fields generated by the surrounding nuclear dipoles and the circulating electrons must oscillate at the appropriate frequency. In the simple case of a one spin system this frequency is termed the Larmor frequency ω_0, specific for the nucleus under consideration. In this context, the types of molecular motion responsible for generating the right range of frequencies to induce relaxation are the rotational diffusion of the molecule, characterised by the rotational correlation time τ_c, and the internal motions within the molecule due to the local fluctuations of the NH bond vector (Keeler, 2005). It is important to note that when studying a two-spin system, in this case the ^{1}H-^{15}N moiety, the dipolar interaction induces transitions between four

distinct spin energy levels that exist under the influence of the static magnetic field B_0. These transitions are all possible pathways for relaxation between specific spin energy states and occur at frequencies different than the ones defined for an isolated spin. In fact, the existence of different energy states implies that the longitudinal, the transverse and the cross-relaxation rates of the ^{15}N nuclei will depend not only upon transitions occurring at the Larmor frequency of both nuclei ω_H and ω_N but will also depend on the motions present at frequencies such as the sum ($\omega_H + \omega_N$) and the difference ($\omega_H - \omega_N$) of the Larmor frequencies. In the particular case of the transverse relaxation rate, it will also be dependent upon molecular motions present at zero frequency. However, the origin of these is beyond the scope of this thesis.

^{15}N R_1 and R_2 and {^1H}-^{15}N NOE parameters are functions of the spectral density function

In order to define the equations describing the ^{15}N R_1 and R_2 and {^1H}-^{15}N NOE parameters, it is necessary to introduce the concept of the spectral density function $J(\omega)$. Simplistically, this function effectively reports on the amount of magnetic field fluctuations present in the sample as a function of the frequency ω. Since R_1, R_2 and {^1H}-^{15}N NOE relaxation parameters are sensitive to these field fluctuations, they can be cast as functions of $J(\omega)$ at the five different frequencies described above. These equations account for relaxation processes driven by the dipole-dipole interaction and the CSA of the ^{15}N nucleus.

$$R_1 = (d^2/4)[J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N)] + c^2 J(\omega_N) \qquad \text{Equation 5.1}$$

$$R_2 = (d^2/8)[4J(0) + J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)] + (c^2/6)[3J(\omega_N) + 4J(0)] \qquad \text{Equation 5.2}$$

$$\text{NOE} = 1 + (\gamma_H + \gamma_N) d^2 [6J(\omega_H + \omega_N) - J(\omega_H - \omega_N)] T_1 \qquad \text{Equation 5.3}$$

where $T_1 = 1/R_1$.

where d and c are constants defined as:

$$d = (\mu_0 \gamma_H \gamma_N h / 8\pi^2) \langle r_{NH}^{-3} \rangle \qquad \text{Equation 5.4}$$

$$c = \Delta\sigma \omega_N / \sqrt{3} \qquad \text{Equation 5.5}$$

where μ_0 is the permeability constant ($4\pi \times 10^{-7}$ H m^{-1}), h is the Planck's constant, γ_H and γ_N are the gyromagnetic ratios of the hydrogen and nitrogen atoms respectively, r_{NH} is the length of the N-H internuclear vector, $\Delta\sigma$ is the CSA of the ^{15}N spin, assumed to be axially symmetric, and ω_H and ω_N are the Larmor frequencies of the ^{1}H and ^{15}N nuclei respectively.

However, the values of the spectral density function at the five frequencies described can not be obtained from the measurement of only three relaxation parameters. An assumption of the form of this function is necessary. The most common approach has been to assume that the dynamics of the amide bond vector derive from the product of two exponentially decaying correlation functions, one describing the overall isotropic tumbling of the molecule characterised by the correlation time τ_c, and the second describing the internal motion of the bond vector characterised by the effective correlation time τ_e and by a generalized order parameter S^2 that reports on the spatial restriction of each internuclear vector (Equation 5.6),

$$J(\omega) = \frac{S^2\tau_c}{(1+\omega^2\tau_c)} + \frac{(1-S^2)\tau}{(1+\omega^2\tau^2)} \qquad \text{Equation 5.6}$$

where $\tau = \tau_c \tau_e / (\tau_c + \tau_e)$. This approach has been termed the "model-free" formalism of Lipari and Szabo (Lipari and Szabo, 1982a,b). With this treatment, there is no specific assumption on the nature of the internal motion of the NH bond vector, as was the case for other methods. This approach was later extended by Clore and co-workers to account for cases in which a third time constant τ_s for slower motions is needed to describe the data (Clore *et al.*, 1990).

An alternative approach to the analysis of NMR relaxation data of a protein has come from the need to abolish effectively the premise of the form of the spectral density function. Peng and Wagner first proposed the spectral density mapping procedure (Peng and Wagner, 1992). Motional information can be directly extracted from the relaxation parameters by exploiting the fact that these parameters can be expressed as a weighted sums of $J(\omega)$ at various different frequencies. Rather than trying to estimate the entire functional form of $J(\omega)$, motional information could be sampled by

calculating $J(\omega)$ at 0, ω_N, ω_H, $\omega_H \pm \omega_N$ frequencies which are ultimately the frequencies that contribute to the relaxation process. In this vein, the authors have extracted additional relaxation parameters such as the longitudinal two-spin order relaxation rate $R_{NH}(H_zN_z)$, the anti-phase transverse relaxation rate $R_{NH}(H_zN_{x,y})$ and the amide proton longitudinal relaxation rate $R_H(H_z)$ in order to allow for the exact determination of the values of $J(\omega)$ at all five frequencies. However, there are several limitations to this approach, particularly with the assumption of the form of the relaxation decay rate of the additional parameters measured. In fact an exponential decay fails to describe correctly the relaxation of the H_zN_z, $H_zN_{x,y}$ and H_z magnetisations. Moreover, the method relies on the differences between relaxation rates of similar magnitude which renders the analysis more prone to experimental errors.

To circumvent this apparent limitation, Farrow *et al.* (1995) proposed some general assumptions which led to the simplification of the functional form of $J(\omega)$. In this respect, the high frequency spectral density terms are of approximately equal magnitude *i.e.* $J(\omega_H \pm \omega_N) \approx J(\omega_H)$. With this approximation the spectral density can now be sampled at three different frequencies which effectively report on motions at the pico- to nanosecond timescale and the slower motions at the micro- to millisecond timescale. This simpler treatment is termed the reduced spectral density mapping.

Overview of motional timescales characterised by NMR

The two protocols proposed above for the analysis of the ^{15}N relaxation parameters have proven successful in the first order analysis of protein molecular dynamics. The earliest study able to show the potential of ^{15}N dynamics focused on staphylococcal nuclease (Kay *et al.*, 1989). The results clearly showed the propensity of loops to demonstrate enhanced mobility in comparison with that of regular secondary structure elements which remained rather rigid, a view that improved our perception of available X-ray structures. Since then, many experimental advances have increased the reliability of the pulse sequences employed in the characterisation of protein dynamics. The fast motions in the picosecond to nanosecond timescale are commonly detected through the measurement of the ^{15}N R_1 and R_2 relaxation rates and through the steady state $\{^1H\}$-^{15}N NOE parameter (Figure 5.1). These motions are generally sampled by flexible loops in proteins, by the rotation of side chains in residues or by the overall tumbling of the molecule. The slower motions in the milli- to microsecond timescale

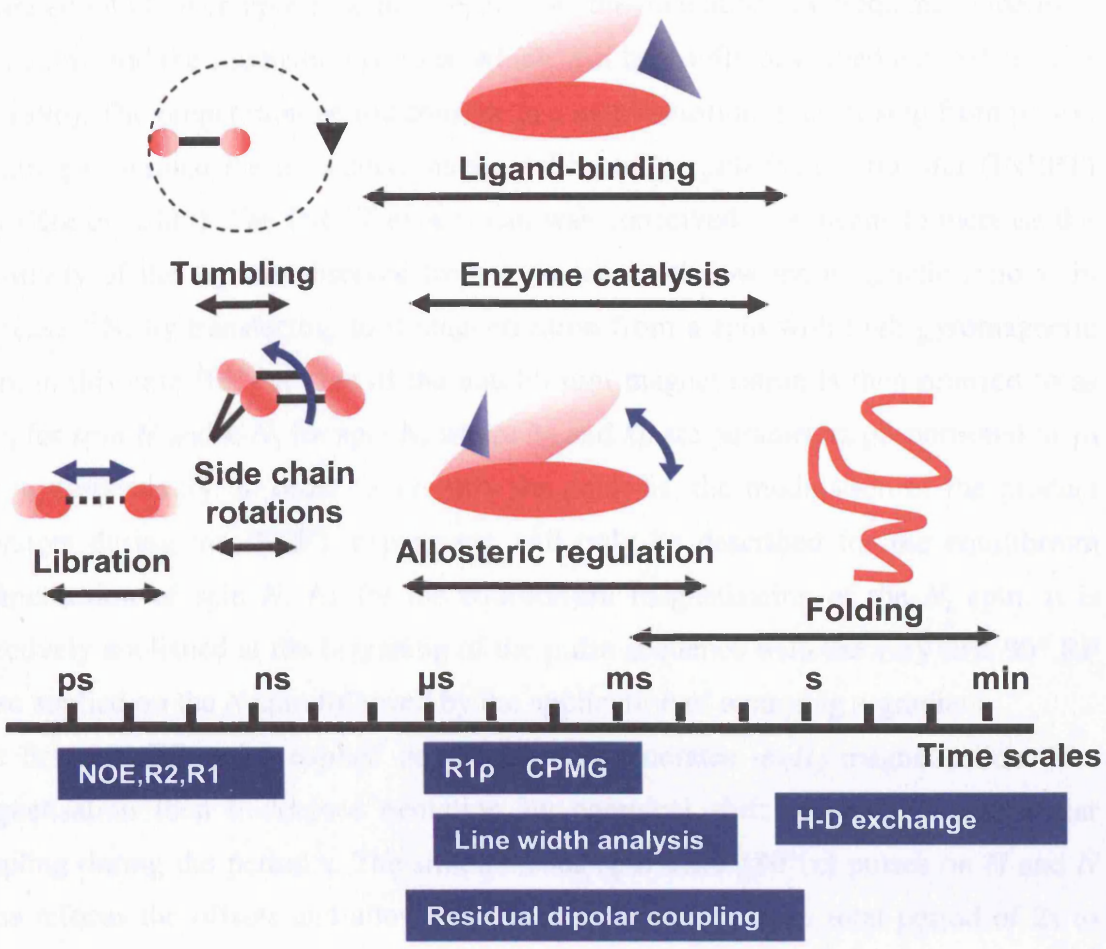

Figure 5.1. Approximate time ranges of biological protein motions (top) and the NMR parameters which can report on them (bottom, see text).

Δ-180°-Δ-acquire). The first 90° RF pulse generates transverse magnetisation which decays during the period Δ. The 180° RF reverses the sense of the precession of the spins and refocuses the evolution due to the chemical shift offsets and B_0 field inhomogeneity.

The NMR pulse sequences used for the measurement of the ^{15}N R_1 and R_2 relaxation rates of *Pa*DDAH are shown in Figure 5.2. Generally, ^{15}N relaxation experiments are separated into five components, the preparation, the relaxation, the frequency labelling, the mixing and the acquisition periods, which will be briefly described here (Palmer *et al.*, 1996). The preparation period consists in a magnetisation transfer step from proton to nitrogen, termed the insensitive nuclei enhanced by polarisation transfer (INEPT) step (Keeler, 2005). The INEPT experiment was conceived as a means to increase the sensitivity of the signals observed from a nucleus with low gyromagnetic ratio γ, in this case ^{15}N, by transferring to it magnetisation from a spin with high gyromagnetic ratio, in this case ^{1}H. The size of the equilibrium magnetisation is then referred to as $k_H H_z$ for spin H and $k_N N_z$ for spin N, where k_H and k_N are parameters proportional to γ_H and γ_N respectively. In order to simplify the analysis, the modulation of the product operators during the INEPT experiment will only be described for the equilibrium magnetisation of spin H. As for the equilibrium magnetisation of the N_z spin, it is effectively abolished at the beginning of the pulse sequence with the very first 90° RF pulse applied on the N spin followed by the application of a purging z-gradient.

The first 90° RF pulse applied on the spin H generates $-k_H H_y$ magnetisation. The magnetisation then undergoes evolution by chemical shift offset (Ω_H) and scalar coupling during the period τ. The simultaneous spin echo 180°(x) pulses on H and N spins refocus the offsets and allow the coupling to evolve for a total period of 2τ to generate the antiphase term $-k_H 2H_x N_z$. The following two RF 90° pulses along the y-axis on H and along the x-axis on N transfer the anti-phase state created on H to the spin N, generating the term $-k_H 2H_z N_y$. At this stage, the pulse sequence presented here diverges from the general scheme described in Palmer *et al.* (1996). The second part of the R_1 and R_2 relaxation experiments generally consists in the relaxation component, however in these experiments the frequency labelling period has been incorporated prior to the relaxation delay. During the frequency-labelling period, the chemical shifts due to the N spin are recorded during the t_1 delay increments to generate the indirect dimension of the two dimensional (2D) spectra. By mapping the chemical shift evolution of the ^{15}N spin before relaxation, any possible frequency due to chemical

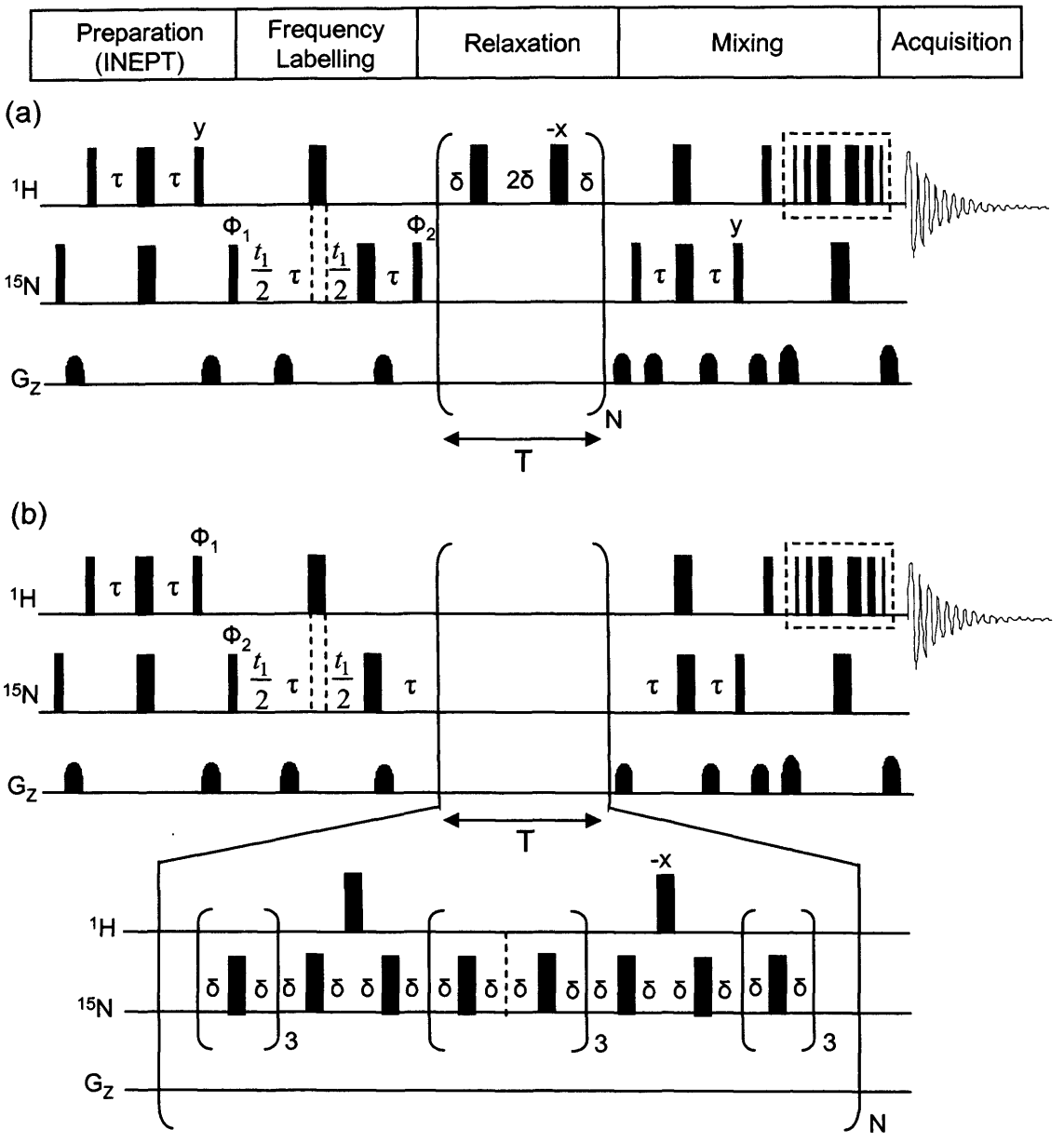

Figure 5.2. Pulse sequences used for the estimation of ^{15}N (a) R_1, (b) R_2 spin relaxation parameters of R40E-R98H *Pa*DDAH incorporating WATERGATE (Piotto *et al.*, 1992) solvent suppression (enclosed in the dashed box). The experiments are commonly separated into five periods indicated in top of the figure. In all sequences narrow and wide pulses indicate 90° and 180° pulses, respectively. All pulses are applied along the x-axis unless indicated otherwise. Z-gradient pulses are grey. Black gradient pulses are the encoding and decoding gradients for solvent suppression. The phase cycling used in the R_1 experiment was Φ_1 = x,-x; Φ_2 = y,-y, in the R_2 experiment Φ_1 = y,-y Φ_2 = x,-x. The receiver phase cycling for both experiments was x,-x,-x,x.

exchange is also being mapped. Hence, the relaxation rates for both slow exchanging peaks can be estimated independently in the subsequent steps. Evidently this feature is only useful in cases where definite chemical exchange occurs for a particular ^{15}N spin, however it has been left in R_1 and R_2 relaxation pulse sequences used at the NIMR.

The INEPT transfer sequence continues with the application of two other sequential RF 180°(x) pulses on both nuclei. The offsets are refocused but the coupling evolves for an extra period of 2τ, at which stage the anti-phase term $-k_H 2H_z N_y$ is converted to $k_H N_x$. This additional [-τ-180°H,N-τ-] element creates in-phase ^{15}N magnetisation in the transverse plane. This sequence has been referred to as refocused-INEPT in which the in-phase signal from the N spin is proportional to k_H and not k_N so that it is only derived from transferred magnetisation.

In R_1 relaxation experiments (Figure 5.2a), an additional 90° RF pulse is applied to spin N along the y-axis which transfers the in-phase transverse coherence $k_H N_x$ to the z-axis. In R_2 experiments this pulse is not necessary since the magnetisation is already in the transverse plane prior to the relaxation period.

The third part of the experiments consists of the relaxation period T. The delay T is increased by repetition in a time series of two dimensional (2D) NMR spectra to generate the signal decay curve. During this period, the two 180° pulses applied to the H spin effectively abolish cross correlation effects between ^1H-^{15}N dipole-dipole and CSA relaxation mechanisms by interchanging the ^1H spin state which averages the relaxation rates of each of the multiplet components suppressing the interference between relaxation mechanisms (Kay *et al.*, 1992b). In the case of the R_2 experiment, the proton 180° pulses are applied at an even echo midway though the Carr-Purcell-Meiboom-Gill (CPMG) pulse train $[\delta\text{-}180°\text{-}\delta]_3$ (Figure 5.2b). The CPMG pulse train applied to the nitrogen spin is designed to refocus the resonance offset, field inhomogeneity, chemical exchange and scalar relaxation (Palmer *et al.*, 1992). In addition, in-phase magnetisation of the ^{15}N nuclei, N_y, left in the transverse plane during the relaxation delay is prone to evolve as a result of scalar coupling to generate the anti-phase term $-2H_z N_x$. The presence of proton anti-phase magnetisation can lead to additional proton-proton dipolar couplings which are another contributor to relaxation. It is important then to minimise the evolution of the NH scalar coupling by setting the delay period δ in the CMPG pulse train shorter than the inverse of the ^1H-^{15}N coupling constant J_{NH}.

During the mixing period, the appropriate ^{15}N coherence is transferred to the *H* spin into observable signals through a reverse INEPT sequence. Water suppression is achieved with the WATERGATE sequence (Piotto *et al.*, 1992). Finally, in the acquisition period the proton transverse magnetisation is recorded to form the directly acquired dimension of the 2D spectrum.

A final consideration regards the application of pulsed field gradients within the pulse sequences. These gradients are designed to suppress the imperfections arising from pulses when inserted on opposite sides of simultaneous $^{1}H/^{15}N$ 180° pulses. They ensure that only transverse magnetisation present before and after the application of the pulse pair is refocused. Similarly, they act as a z-filter when the magnetisation is along z in order to aid in the suppression of the water resonance. The final two gradients in the pulse sequence are the encoding and decoding gradients part of the WATERGATE pulse scheme to also enable water suppression.

The NMR pulse sequence used for the measurement of the steady-state heteronuclear ^{15}N NOE values of R40E-R98H *Pa*DDAH is shown in Figure 5.3. The purpose of this experiment is to observe the maximum "enhancement" of magnetisation of the ^{15}N nucleus as the ^{1}H spin is saturated. This parameter is measured with two separate 2D heteronuclear correlation experiments. One experiment includes a presaturation period consisting of a train of 180° pulses applied the ^{1}H spins, the other experiment omits this period. The intensity of the ^{15}N magnetisation can then be compared in the two spectra to determine the relative enhancement due to ^{1}H saturation. The initial INEPT sequence used in the R_1 and R_2 experiments is also omitted in NOE experiments; consequently ^{15}N magnetisation excited by the first ^{15}N 90° pulse is directly transferred to the proton for detection through a reverse refocused INEPT pulse scheme. Since the magnetisation acquired derives solely from the magnetisation of the ^{15}N spin, heteronuclear NOE experiments are generally less sensitive then ^{15}N relaxation experiments.

Extracting ^{15}N R_1 and R_2 relaxation rates and $\{^{1}H\}$-^{15}N heteronuclear NOE values

^{15}N R_1 and R_2 rates were estimated for the monomeric *Pa*DDAH at 26°C and at ^{1}H NMR frequencies of 500, 600 and 800 MHz. The signal intensities of well resolved cross peaks were extracted from each NH correlation spectrum at a particular relaxation delay following the protocol described in Chapter 2, Section 2.4. The analysis included 197 amide cross peaks obtained at 500 and 600 MHz and 224 amide

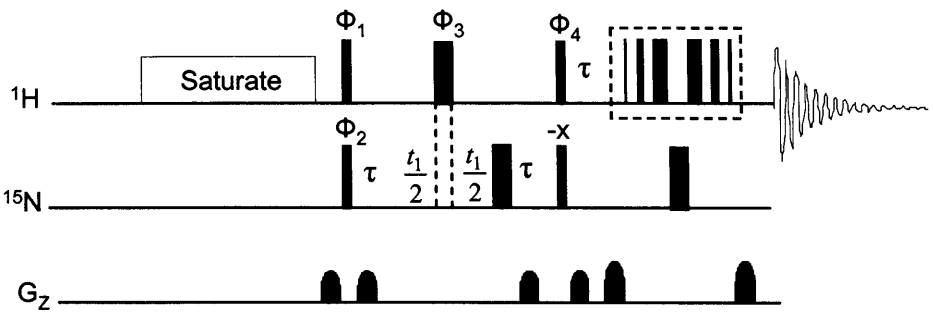

Figure 5.3. Pulse sequence used for the estimation of $\{^1H\}$-^{15}N heteronuclear NOE spin relaxation parameter of R40E-R98H *Pa*DDAH incorporating WATERGATE (Piotto *et al.*, 1992) solvent suppression (enclosed in the dashed box). Narrow and wide pulses indicate 90° and 180° pulses, respectively. All pulses are applied along the x-axis unless indicated otherwise. Z-gradient pulses are grey. Black gradient pulses are the encoding and decoding gradients for solvent suppression. The phase cycling used in the NOE experiment was Φ_1 = 2(x), 2(y), 2(-x), 2(-y); Φ_2 = x,-x; Φ_3 = 2(y), 2(-x), 2(-y), 2(x); Φ_4 = 2(-x), 2(-y), 2(x), 2(y). The receiver phase cycling for both experiments was x,-x,y,-y,-x,x,-y,y.

cross peaks at 800 MHz (out of 241 possible). In both ^{15}N R_1 and R_2 experiments, the signal intensity of the 2D cross peaks is a function of the duration of the delay period, t (Equation 5.7).

$$I(t) = I_0 \exp\left[\frac{-t}{T_i}\right] \qquad \text{Equation 5.7}$$

where $I(t)$ is the measured signal intensity at time t, I_0 is the initial signal intensity at t=0, t is the delay period and T_i is the relaxation time where i = 1 or 2 and $R_i = 1/T_i$. The measured signal intensities were then fitted to Equation 5.7 (two parameter fit) and the relaxation rates R_i were calculated using *Mathematica*® software. Representative fits of the data to extract R_1 and R_2 rate constants obtained at ^1H frequency of 600 MHz are shown in Figure 5.4 for several residues. The intensities have been plotted as a function of relaxation delay and the repeated time points are also shown. The percentages of error estimated from the exponential fit to the original data for ^{15}N R_2 rate constants after Monte Carlo simulations were 4.2, 3.4 and 2.7% at 500, 600 and 800 MHz respectively. The estimated errors for the ^{15}N R_1 experiments were slightly lower at 2.7, 2.9 and 1.4% for 500, 600 and 800 data sets respectively.

Similarly, the steady-state NOE values for R40E-R98H *Pa*DDAH have been estimated following the protocol described in Chapter 2, Section 2.4. Since the error was estimated from the spectral noise, the data collected at 800 MHz has the lowest uncertainty at 2.8% whereas 500 and 600 MHz NOE values have an average percentage of error around 6.4%. The values for all three relaxation parameters ^{15}N R_1, R_2 and heteronuclear NOE are shown at three different magnetic field strengths as a function of residue number in Figure 5.5 and the values are shown in Tables A.III, A.IV and A.V in Appendix. Initial inspection of the data shows a similar pattern of R_1 rates at the three magnetic fields, the same trend is verified for the R_2 and NOE values (Figure 5.5). As expected the ^{15}N R_1 rates are highest at 500 MHz, the average values of R_1^{500}, R_1^{600} and R_1^{800} being 1.24 ± (0.08) s^{-1}, 0.98 (± 0.09) s^{-1} and 0.67 (±0.07) s^{-1}, respectively. In terms of ^{15}N transverse relaxation rates, theory predicts that the contribution from dipole-dipole relaxation to R_2 is expected to be identical at all fields. However, the contribution of chemical shift anisotropy (CSA) to the R_2 rate is proportional to $(\omega_N)^2$. Consequently, R_2^{800} values should be higher than R_2^{600} and R_2^{500}, a result which was verified experimentally with the average values of R_2^{500}, R_2^{600}, R_2^{800} being 15.18 (±1.47) s^{-1}, 16.98 (±1.67) s^{-1} and 20.22 (±1.64) s^{-1}, respectively. Similarly, the analysis of Equation 5.3 predicts that the $\{^1H\}$-^{15}N NOE

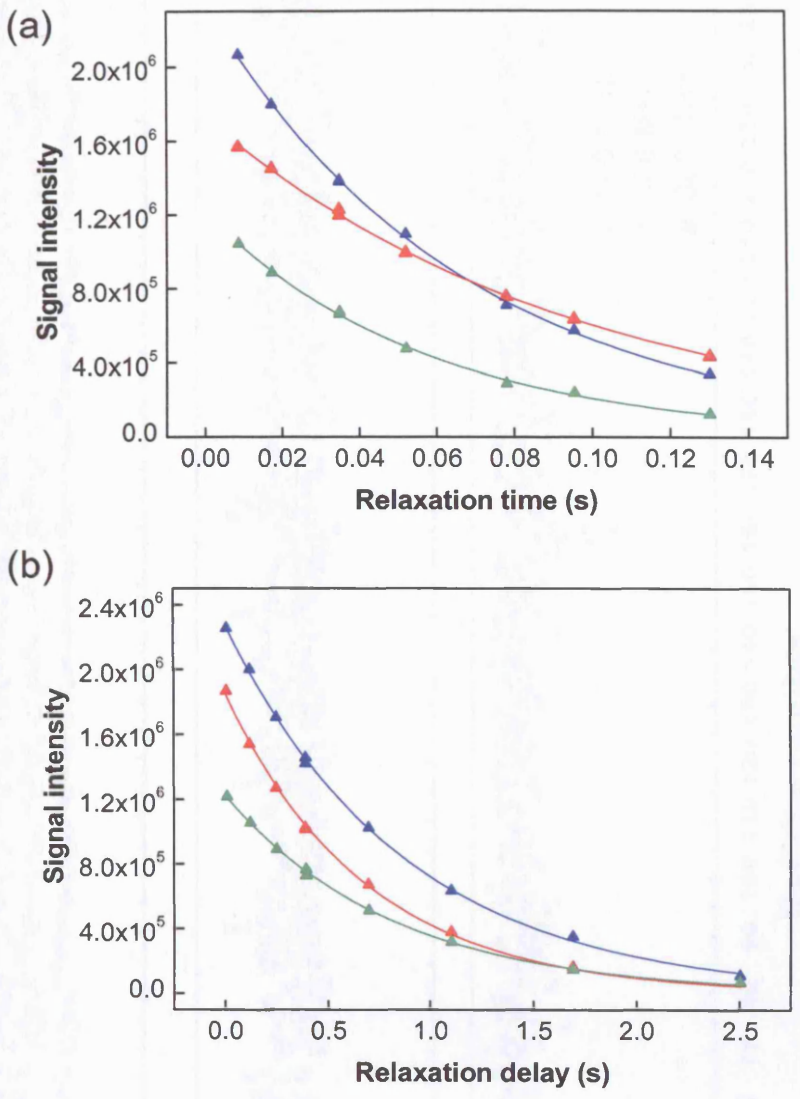

Figure 5.4. Representative single exponential decay fits of the cross peak intensities recorded at ^{1}H frequency of 600 MHz as a function of (a) ^{15}N transverse relaxation delay, (b) longitudinal relaxation delay. The decay curves in blue, red and green are ^{15}N R_2 and R_1 relaxation decays of residues Val 104, Leu 18 and His 162.

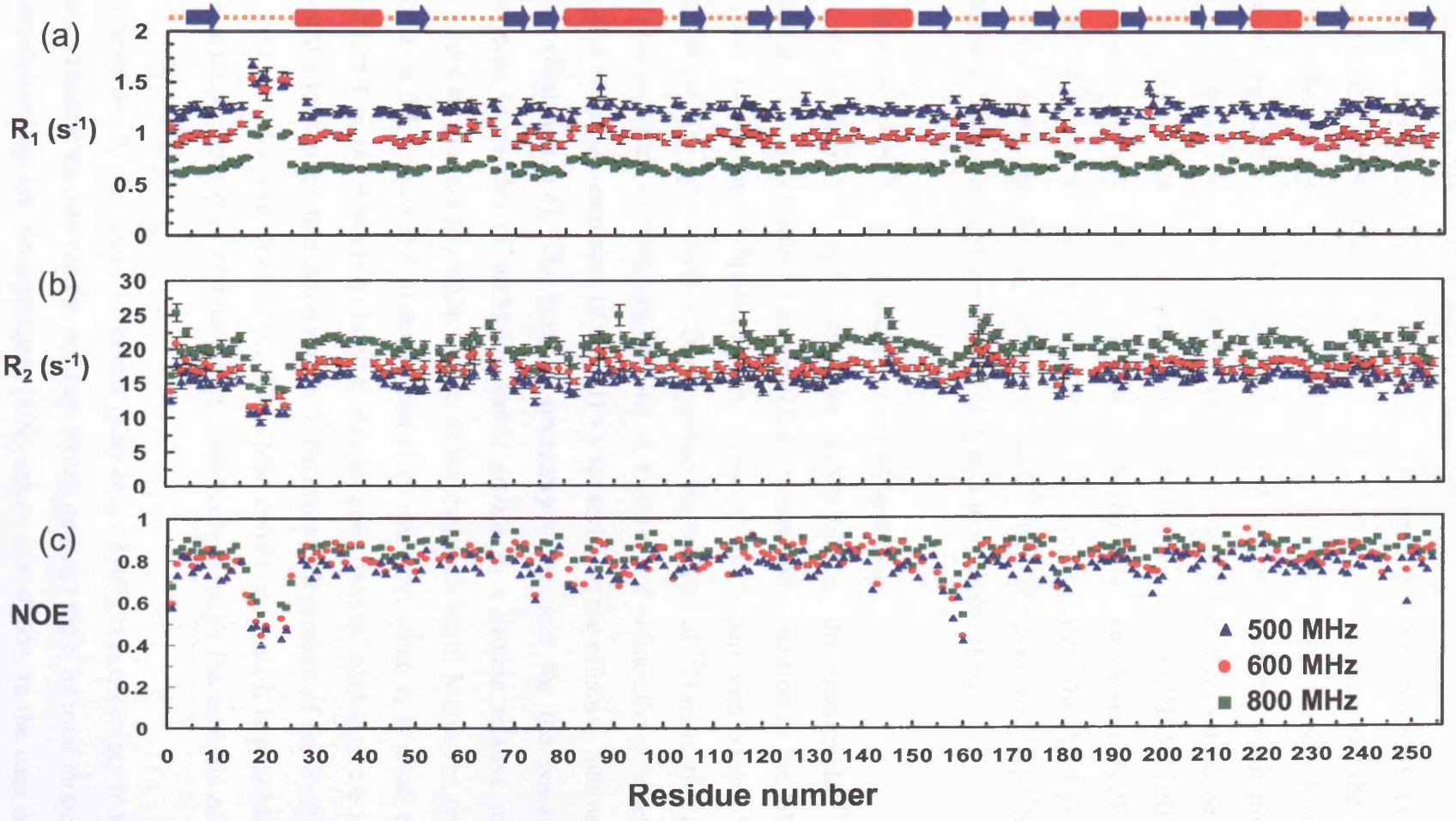

Figure 5.5. (a) Longitudinal ^{15}N relaxation rates R_1, (b) transverse ^{15}N relaxation rates R_2 and (c) $\{^1H\}$-^{15}N heteronuclear NOE values of R40E-R98H *Pa*DDAH extracted at three different magnetic field strengths (blue triangle at ^1H nominal frequency of 500 MHz, red circles at 600 MHz and green squares at 800 MHz) plotted as a function of residue number. Secondary structural elements are indicated at the top of the Figure.The standard deviations of the heteronuclear NOE values are omitted for clarity.

enhancements increase with increasing magnetic field strengths. This prediction is also confirmed as shown by the average values of NOE^{500}, NOE^{600} and NOE^{800} of 0.75 (±0.09), 0.79 (±0.09), 0.84 (±0.07), respectively.

Taken as a whole the values of R_1 and R_2 rate constants are remarkably consistent over most of the protein chain, which is in line with the expectation that the protein has a stable globular fold. There is mainly one region of the protein for which either R_1 or R_2 values are significantly different from the mean values over the whole protein (Figure 5.5). R_1 values are higher for residues 17 through 24 and R_2 values decrease in the same region. These residues are located in a loop region of *Pa*DDAH believed to close down on the active site upon ligand binding. The steady-state NOE values also decrease around the same loop region. In addition, residues Glu 158, Lys 159 and Val 160 also have significantly lower NOE values than the average $\{^1H\}$-^{15}N NOE value obtained. These residues are also located on a loop region (L6).

Estimation of the isotropic rotational correlation time

As discussed in Section 5.1, the relationship between the measurable ^{15}N relaxation parameters in a two spin system and the reorientational motion of the NH vector in a protein is commonly interpreted in terms of the Lipari and Szabo "model-free" formalism. The application of this model to the analysis of ^{15}N relaxation data assumes that the molecule tumbles isotropically in solution and defines the parameters S^2 as the degree of spatial restriction of the NH vector and τ_e as the effective internal correlation time (Equation 5.6). The latter is necessary to account for the possible differing motional timescales of backbone amide groups in a protein relative to the overall isotropic rotational correlation time τ_c of the molecule itself. Moreover, the analysis of Equation 5.6 reveals that under certain circumstances, when τ_e is small (i.e <100 ps) and $\tau_c > 1$ ns (*i.e* when both internal motion and chemical exchange are not dominant contributions to the relaxation rates), S^2 becomes independent of the R_2/R_1 ratio which in turn becomes a function of τ_c. With these conditions met, it is possible to estimate the isotropic rotational correlation of a molecule through the analysis of the average R_2/R_1 ratio, $\langle R_2/R_1 \rangle$.

As described by Kay and co-workers (Kay *et al.*, 1989), it is necessary to remove from these calculations data for the residues which undergo rapid internal motion which are characterised by low heteronuclear NOE values (i.e <0.65). In the case of *Pa*DDAH,

the application of this filter excluded 15 residues from the 500 MHz data, 12 residues from the 600 MHz data and 8 residues from the 800 MHz data. Additionally, the contribution of chemical exchange to R_2 rates can lead to an over estimation of $\langle R_2/R_1 \rangle$ and consequently an erroneously long correlation time (Tjandra *et al.*, 1995). Therefore a standard deviation filter is applied to exclude residues for which $\langle R_2/R_1 \rangle$ is significantly different from the sample mean (Equation 5.8), *i.e.* when,

$$\left| \frac{\langle R_1 \rangle - R_{1,i}}{\langle R_1 \rangle} - \frac{\langle R_2 \rangle - R_{2,i}}{\langle R_2 \rangle} \right| > 1.5 \times \text{SD} \qquad \text{Equation 5.8}$$

where $\langle R_1 \rangle$ and $\langle R_2 \rangle$ are the sample average longitudinal and transverse relaxation rates respectively, $R_{1,i}$ and $R_{2,i}$ are the individual NH group relaxation rates and SD is the standard deviation over the complete data set of unfiltered ^{15}N R_2/R_1 values (Tjandra *et al.*, 1995). This filter excluded 5 additional residues from the 500 MHz data, 7 residues from the 600 MHz and 8 residues from the 800 MHz data set.

With a suitable data set filtered as described above, τ_c can be calculated by a minimization of Equation 5.9

$$\chi^2 = \left[\frac{R_2(\tau_c)}{R_1(\tau_c)} - \langle R_2 / R_1 \rangle \right]^2 \qquad \text{Equation 5.9}$$

where $R_1(\tau_c)$ and $R_2(\tau_c)$ are functions describing the effect of molecular tumbling frequency on the respective relaxation rates and $\langle R_2/R_1 \rangle$ is the average experimental rate ratio.

The average $\langle R_2/R_1 \rangle$ ratio values obtained in this manner at three different magnetic fields are R_2^{500}/R_1^{500} of 12.55 (±0.68), R_2^{600}/R_1^{600} of 17.89 (±1.00) and R_2^{800}/R_1^{800} of 30.54 (±1.97) (Table 5.1; Figure 5.6). Subsequently, the values of τ_c obtained at 500, 600 and 800 MHz after performing Monte Carlo simulations are 13.14 (±0.16) ns, 13.24 (±0.16) ns and 13.14 (±0.19) ns, respectively. These values correlate well at different magnetic field strengths and are in the proper range for a 29 kD protein (Table 5.1). The mean value 13.17 (±0.13) ns was then taken as a conservative estimate of the isotropic rotational correlation time of R40E-R98H *Pa*DDAH at 26°C.

In order to address the quality of the data sets acquired at different fields, theoretical calculations were performed to derive R_2 and R_1 rates in which S^2 was set to 0.85, τ_e to 0 and τ_c to 13.17 ns in Equation 5.3. The consistency of the ^{15}N parameters obtained at 500, 600 and 800 MHz was estimated in terms of the theoretical values predicted for

Table 5.1 Average ^{15}N heteronuclear spin relaxation parameters for R40E-R98H *Pa*DDAH recorded at 26°C at different magnetic field strengths

	500 MHz		600 MHz		800 MHz	
Parameter	All[a]	Filtered[b]	All[a]	Filtered[b]	All[c]	Filtered[d]
$^{15}N\ \langle R_1 \rangle\ (s^{-1})$	1.24 ± 0.08	1.23 ± 0.05	0.98 ± 0.09	0.96 ± 0.05	0.67 ± 0.07	0.66 ± 0.04
$^{15}N\ \langle R_2 \rangle\ (s^{-1})$	15.18 ± 1.47	15.35 ± 0.86	16.98 ± 1.67	17.18 ± 0.92	20.23 ± 1.64	20.26 ± 1.19
$\langle R_2/R_1 \rangle$	-	12.55 ± 0.68	-	17.89 ± 1.00	-	30.54 ± 1.97
$\{^1H\}$-^{15}N <NOE>	0.75 ± 0.09	0.78 ± 0.05	0.79 ± 0.09	0.81 ± 0.05	0.84 ± 0.07	0.85 ± 0.05
τ_c (ns)[e]	13.14 ± 0.16		13.24 ± 0.16		13.14 ± 0.19	

[a]Average values calculated from the data obtained for 199 well resolved cross peaks.
[b]Average values re-estimated using NOE and R_2/R_1 filters (179 cross peaks left in the analysis of the 500 MHz data and 182 residues left in the analysis of the 600 MHz data).
[c]Average values calculated from the data obtained for 226 well resolved cross peaks.
[d]Average values re-estimated using NOE and R_2/R_1 filters (211 cross peaks left in the analysis of the 800 MHz data).
[e]Calculated assuming isotropic rotational diffusion and using the filtered data.

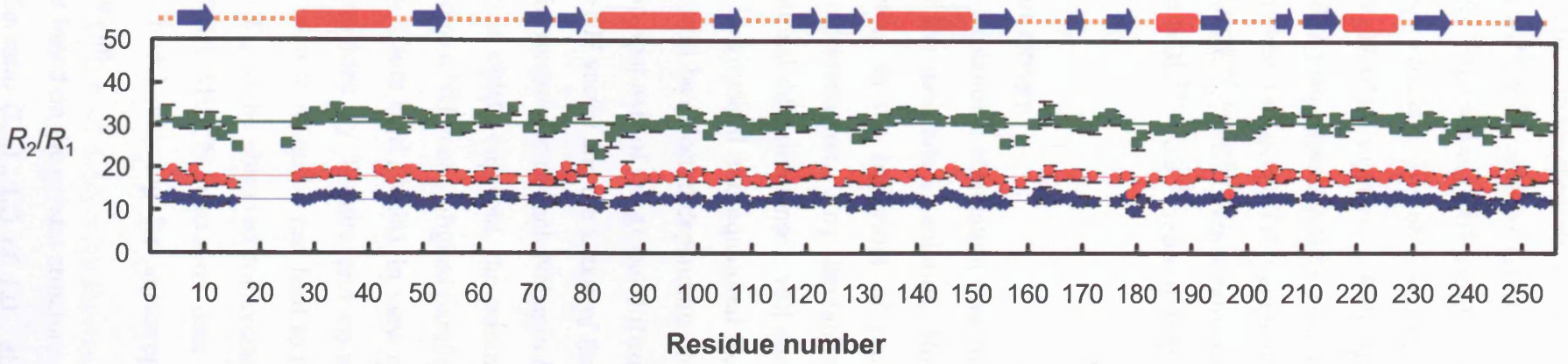

Figure 5.6. ^{15}N R_2/R_1 ratios of R40E-R98H *Pa*DDAH backbone amides plotted as a function of residue number and extracted at three different magnetic field strengths (blue triangle at ^{1}H nominal frequency of 500 MHz, red circles at 600 MHz and green squares at 800 MHz). Secondary structural elements are indicated at the top of the Figure.

R_2^{800}/R_2^{600} and R_1^{800}/R_1^{600} ratios which were 1.17 and 0.68 respectively. The observed 800/600 ratios agree well with the theoretical ones with an average value of R_2^{800}/R_2^{600} of 1.18 (±0.09) and R_1^{800}/R_1^{600} of 0.69 (±0.05). Similarly, the predicted values of R_2^{600}/R_2^{500} and R_1^{600}/R_1^{500} ratios were 1.07 and 0.76 respectively and were compared to the experimental values obtained. The observed R_2^{600}/R_2^{500} and R_1^{600}/R_1^{500} ratios were 1.11 (±0.08) and 0.79 (±0.05) respectively. Although both values of the experimental ratios are slightly higher than expected, they remain within experimental uncertainty. The consistency of the heteronuclear NOE values were also verified as the theoretical NOE^{800}/NOE^{600} and NOE^{600}/NOE^{500} ratios estimated at 1.04 and 1.02 respectively, matched the observed NOE^{800}/NOE^{600} ratio of 1.05 (±0.09) and NOE^{600}/NOE^{500} ratio of 1.02 (±0.09).

Rotational diffusion anisotropy

The determination of a rotational correlation time for R40E-R98H *Pa*DDAH assumed isotropic tumbling of the molecule in solution. However, such an approximation is generally a shortcoming as the behaviour of proteins tends to deviate from the idealistic spherical representation. Any deviation from a sphere, described as anisotropy of the rotational diffusion tensor, will equate to different rates of rotational diffusion around the longitudinal and equatorial axes of the protein. The relaxation rates of the NH spins will be affected depending on the orientation of the NH vector with respect to the principal axis of the global diffusion tensor. In particular, the angle θ formed between the NH vector and the z-axis of the diffusion tensor D_{zz} is critical in the determination of the magnitude of each NH spin R_2/R_1 ratio. It has been shown that for an axially symmetric oblate ellipsoid, the minimum and maximum experimental R_2/R_1 ratios correspond to a NH vector aligned parallel, *i.e.* $\theta = 0°$ or perpendicular, $\theta = 90°$ to D_{zz} respectively (Clore *et al.*, 1998). In view of these considerations and on the additional evidence provided by Tjandra and co-workers demonstrating that even slight anisotropic diffusion in ubiquitin can lead to the artifactual identification of an exchange contribution R_{ex} to the observed transverse relaxation rates (Tjandra *et al.*, 1995), R40E-R98H *Pa*DDAH ^{15}N relaxation data sets were further scrutinised to determine whether a model allowing for anisotropic diffusion best described the enzyme rotational diffusion.

The initial calculations based on the crystal structure of *Pa*DDAH yield an estimate of the moments of inertia ratio ($I_{xx}:I_{yy}:I_{zz}$) of 1.0 : 0.97 : 0.39 for the homodimeric

*Pa*DDAH; the corresponding ratio predicted for the isolated protomer is 1.0 : 0.86 : 0.80. Based upon this assessment the homodimer would be expected to show diffusional characteristics consistent with an oblate ellipsoid with $I_{\perp}/I_{\parallel} \sim 2.5$, whereas the protomer should show only a slight deviation from isotropic tumbling ($I_{\perp}/I_{\parallel} \sim 1.2$). To estimate this deviation, the degree of anisotropy A (Equation 5.10) can be obtained from a least squares optimisation on the basis of the measured ^{15}N R_2/R_1 ratios and the orientation of the internuclear NH vector provided by the crystal structure (Tjandra *et al.*, 1995, 1996).

$$A = \frac{2D_{zz}}{D_{xx} + D_{yy}} \qquad \text{Equation 5.10}$$

where D_{xx}, D_{yy} and D_{zz} are the principal axes of the diffusion tensor. Prior to the analysis, hydrogen atoms were added to the crystal structure of the citrulline-bound *Pa*DDAH with the program X-PLOR (Brünger, 1992). In order to choose the diffusion model that best describes the ^{15}N relaxation data acquired for the monomeric *Pa*DDAH, the programs R2R1_diffusion and quadric_diffusion (Lee *et al.*, 1997) were used in the analysis. R2R1_diffusion uses the approach described in Tjandra *et al.* (1995) to determine the diffusion tensors for an isotropic and an axially symmetric motional regime by comparing the observed and predicted R_2/R_1 values for each selected NH cross peak. Quadric_diffusion uses the quadratic representation approach to fit the experimentally derived ^{15}N relaxation rates to the isotropic, axially symmetric or fully asymmetric motional regime (Lee *et al.*, 1997). For the isotropic model a single rotational correlation time is returned from both programs. In the case of axially symmetric rotational diffusion – either oblate or prolate forms – the parallel, $D_{\parallel} = 2D_{zz}$ and perpendicular, $D_{\perp} = (D_{xx} + D_{yy})$ components of the diffusion tensor and the angles θ and φ are optimized, where θ is the angle between the N-H interatomic vector and the z-axis of the diffusion tensor, and φ the angle between the projection of the N-H vector on the x-y plane and the x-axis. In the case of a totally asymmetric model six parameters D_{zz}, D_{yy} and D_{xx}, and θ, φ and ψ are optimised in which the Euler angles define the orientation of the diffusion tensor coordinates in the chosen molecular frame (this model is only assessed in the quadric_diffusion program). In all cases, the parameters are extracted by minimising the error function characterised by the difference between predicted and experimental R_2/R_1 ratios for each model. The predicted values are derived from Equations 5.1, 5.2 and from the extended version of

Equation 5.6 as determined by Clore *et al.* (1990) to account for rotational anisotropy (*vide infra*). The Monte Carlo simulated χ^2 distribution and *F*-statistics are then calculated in order to address the quality of the fitting procedure and the comparison of models. Only ^{15}N relaxation data recorded at 500 and 600 MHz were used in this assessment. The data sets were filtered in the same manner as that employed in the determination of the isotropic rotational correlation time of the molecule. However, an additional round of data filtering was performed and residues which exhibited a predicted effective diffusion constant with a high χ^2 residual value (χ^2>100) were eliminated from the fitting procedure (approximately 60 residues were removed from the final calculation). This step further ensures that residues with no internal motions are considered (Butterwick *et al.*, 2004). The results are shown in Table 5.2 and are essentially the same at both magnetic field strengths.

Table 5.2. Anisotropic rotational diffusion parameters for monomeric *Pa*DDAH

Parameter	500 MHz[a]	600 MHz[a]
$D_{\parallel}/D_{\perp}$	1.06 (±0.01)	1.05 (±0.01)
D_{iso} (10^7 s^{-1})[b]	1.26 (±0.02)	1.25 (±0.02)
θ (rad)[c]	1.3	1.5
φ (rad)[c]	5.4	5.2
F-value[d]	25.1 ($p < 10^{-4}$)	14.0 ($p < 10^{-4}$)

[a]The analysis was performed with 132 residues at 500 MHz and 135 at 600 MHz. [b]$D_{iso} = 1/6\tau_c$ where D_{iso} is the isotropic diffusion tensor. [c]Euler angles describing the orientation of the components of the diffusion tensor relative to the inertial tensor frame of the pdb file (1h70). [d]Value given when testing the significance of the axially symmetric diffusion model over the isotropic model.

The axial model was preferred over the isotropic model as best describing the data as indicated by the high *F*-values obtained with either ^{15}N relaxation data sets (Table 5.2). The fully anisotropic model did not improve the fitting of the data. When comparing the χ^2 values derived from the fitting of the data sets to an axially symmetric and a fully asymmetric rotational diffusion models, the program quadric_diffusion estimated an *F*-value of 1.8 (p = 0.16) at 500 MHz and an *F*-value of 0.79 (p = 0.45) at 600 MHz. These values are not statistically significant and do not justify the adoption of the fully asymmetric model.

The data acquired at 800 MHz was in general agreement with these results however the filtering of the data is further complicated by exchange contributions that scale quadratically with the external field strengths. As a result, the χ^2 residual values

obtained from the fitting process of each individual spin to a predicted diffusion constant were much higher than for the data sets recorded at the two other magnetic field strengths. For this reason, the diffusion tensor obtained at 500 and 600 MHz was used in the ensuing model-free analysis.

Although both 500 and 600 MHz ^{15}N relaxation data sets of the monomeric *Pa*DDAH predict a slight deviation from isotropic tumbling and agree well in the determination of the θ and φ angles, the overall accuracy of estimating the global rotational diffusion parameters remains limited. The number of residues involved in the analysis is considerably lower than the total number of residues in the protein, flexible regions are generally absent from the assessment and regions with large R_2/R_1 ratios are also excluded. These approximations are liable to an underestimation of the overall anisotropy of the molecule. In the particular case of *Pa*DDAH, a last shortcoming in the estimation of the diffusion tensor of the molecule is the fact that the pdb file used in the assessment is that of the citrulline-bound state of the enzyme. To evaluate the overall rotation of the apo-form of the molecule, one has to assume that it will not be affected by the ligand present in the active site region. However, it is unlikely that this assumption holds true for the entire protein. The conformation of the L1 loop in particular for which no electron density was found in the X-ray diffraction data of the apo-protein (Murray-Rust *et al.*, 2001) is affected upon binding of the substrate.

5.3 Backbone dynamics of the monomeric *Pa*DDAH: Model-free analysis

The model-free protocol

Having established the degree of rotational anisotropy of *Pa*DDAH, it is now possible to determine the amplitudes and timescales of the intramolecular motions of the enzyme backbone amides. The experimentally accessible ^{15}N relaxation rates R_1, R_2, and NOE values can be expressed as linear combinations of the spectral density function $J(\omega)$, which takes the form presented in Equation 5.11 in the extended model-free formalism of Lipari-Szabo (Clore *et al.*, 1990). This model is formally equivalent to taking into account axially symmetric rotational diffusion anisotropy.

$$J(\omega) = \frac{2}{5} S_f^2 \sum_{j=1}^{3} A_j \left[\frac{S^2 \tau_j}{1 + (\omega \tau_j)^2} + \frac{(1 - S_f^2)\tau_j'}{1 + (\omega \tau_j')^2} \right] \qquad \text{Equation 5.11}$$

where $\tau_j' = \tau_j\tau_e/(\tau_j + \tau_e)$, τ'_j is the effective correlation time dependent on τ_e which in turn is the effective correlation time for internal motion; τ_1, τ_2 and τ_3 are time constants defined as $\tau_1^{-1} = 6D_\perp$, $\tau_2^{-1} = 5D_\perp + D_\parallel$, $\tau_3^{-1} = 2D_\perp + 4D_\parallel$ respectively; the terms A_1, A_2 and A_3 are given by $A_1 = (3\cos^2\theta - 1)^2/4$, $A_2 = 3\sin^2\theta\cos^2\theta$, $A_3 = (3/4)\sin^4\theta$, where θ is the angle between the N-H bond vector and the principal axis of frame of the diffusion tensor; $S^2 = S_f^2 S_s^2$ is the square of the generalised order parameter characterising the amplitude of internal motions, and S_f^2 and S_s^2 are the squares of the order parameters for the fast and slow timescales, respectively. The order parameter S^2 derived for each individual N-H vector in the backbone quantifies the amplitude of motion of this bond. The S^2 values satisfy the inequality $0 \leq S^2 \leq 1$; lower values indicate larger amplitudes of internal motion, revealing regions in the backbone of high flexibility and dynamics on a ps-ns timescale.

Alternatively, conformational exchange processes are represented by an *ad-hoc* extra parameter R_{ex} that refers to dynamics on a μs-ms timescale and is added to the expression defining the transverse relaxation rate, $R_2 = R_2^0 + R_{ex}$ where R_2^0 is the relaxation rate in the absence of exchange.

A complete model-free analysis of internal motion of the monomeric *Pa*DDAH was performed using the Modelfree program (version 4.16) (Palmer *et al.,* 1991b) with ^{15}N spin relaxation data acquired at 500, 600 and 800 MHz. The data sets obtained for each individual NH spin were fit independently to five different motional regimes under the assumption of an axially symmetric diffusion tensor (Mandel *et al.* 1995). Each motional model contained the following adjustable parameters: S^2 (model 1); S^2 and τ_e (model 2); S^2 and R_{ex} (model 3); S^2, τ_e and R_{ex} (model 4); and S_f^2, S_s^2 and τ_e (model 5). Furthermore, the average ^1H-^{15}N distance r_{NH} was assumed to be 1.02 Å and the ^{15}N magnitude of the chemical shift anisotropy (CSA) tensor was set to -170 ppm. The rationale for choosing a larger value of CSA over the commonly used -160 ppm is related to the potential uncertainty over the magnitude of the constant contribution of CSA to the measured R_1 rates and to lesser extent R_2 rates (Tjandra *et al.*, 1996; Kroenke *et al.*, 1999). This contribution varies according to the environment in which the amide ^{15}N spin is located (Sitkoff and Case, 1998), but more importantly in our case, it is larger at the higher magnetic field strengths used in this study (>14.1 T).

The best model for each NH group was determined as described by Mandel *et al.* (1995). The goodness of the fits of the experimental relaxation data to each motional model was estimated by the comparison of the χ^2 function at 95% confidence level of the χ^2 probability distribution determined from 500 Monte Carlo simulations. To assign spins to models 2 and 3, *F*-testing also had to be satisfied at the critical value of α = 0.2. When assigning residues to models 4 and 5, the sum-squared error of the fitting procedure for the spin under consideration had to be equal to 0.

Optimised diffusion tensor and internal motion parameters

For each data set evaluated in Modelfree, the optimised parameters for the diffusion tensor are shown in Table 5.3.

Table 5.3. Optimised anisotropic rotational diffusion parameters of the monomeric *Pa*DDAH after Modelfree analysis

	500 MHz[a]	600 MHz[a]	800 MHz[a]
τ_c (ns)	13.16 (±0.02)	13.27 (±0.02)	13.14 (±0.01)
$D_{\|\|}/D_{\perp}$	1.06 (±0.01)	1.06 (±0.01)	1.07 (±0.01)
θ (rad)[a]	1.6 (±0.1)	1.4 (±0.1)	1.9 (±0.1)
φ (rad)[a]	5.4 (±0.1)	5.2 (±0.1)	5.2 (±0.1)

[a]Euler angles describing the orientation of the components of the diffusion tensor relative to the coordinates of the inertial tensor frame in the pdb file (1h70).

The parameters agree well at all three different magnetic fields apart from a slight increase of the θ angle with the data set acquired at 800 MHz. The model selection profiles for each NH spin are shown in Table 5.4.

Table 5.4. The profiles of model selection for the NH spins analysed[a]

	500 MHz	600 MHz	800 MHz
Model		Number of residues	
1	159	157	168
2	16	13	16
3	7	13	13
4	3	1	6
5	10	11	16

[a]A total of 197 spins were analysed at both 500 and 600 MHz and 224 spins were analysed at 800 MHz. Residues 236 and 249 could not be fit to any motional model with either 500 or 600 MHz relaxation data sets. Residues 13, 15, 16, 46, 236 could not be fit with 800 MHz data.

The vast majority of residues were fit to model 1. Several residues had small τ_e contributions (<50 ps), and the same percentage of residues (~7%) were fit to model 5

with τ_e values in the range of 1 ns. In contrast, few residues exhibited both R_{ex} and τ_e terms as only ~1-3% of residues were fit to model 4. Chemical exchange contributions were observed in 4-6% of residues. Some discrepancies were observed in the case of the active site residue Cys 249 and the residues located prior to the L1 loop. Although Cys 249 was fit to model 1 with the 800 MHz data set, it could not be fit to any motional model with the 500 MHz and 600 MHz data sets. Conversely, residues 13, 15 and 16 were best described by relaxation models 1 (residues 13 and 15) and 5 (for residue 16) with both 500 and 600 MHz relaxation data sets, but these residues could not be fit satisfactorily to any model with the data acquired at 800 MHz.

The internal motion parameters derived from the analysis are plotted as a function of residue number in Figure 5.7. The main conclusion drawn from the inspection of these results is the agreement of the data with regards to the average values of the generalised order parameter $\langle S^2 \rangle$ of the NH bond vectors obtained at three magnetic field strengths. Overall the $\langle S^2 \rangle$ values for the secondary structural elements of the protein estimated at 600 and 800 MHz are essentially the same, 0.87 (±0.07) and 0.87 (±0.06) respectively (Table 5.5) attesting to the rigidity of the entire β/α propeller fold. The $\langle S^2 \rangle$ values estimated at 500 MHz are slightly lower than those obtained at the two other fields, 0.85 (±0.07). Despite this fact, several regions of the protein are identified as inherently flexible by this analysis. These regions display internal motions in the pico- to nano- second timescales as shown in Figure 5.7b. Specifically, residues located in loops L1, L3, L6, residue Gly 179 located in a 3_{10}-helix and residue Asp 244 in loop L7 have significant τ_e contributions. The locations of these loops are illustrated in Figure 5.8a. Loop L1 comprising residues 14-25 has an $\langle S^2 \rangle$ value of ~0.6 and lies over the active site region. Loops L3, L6 and L7 are located at the top entrance of the cavity forming the active site region (Figure 5.8a). Regions of the protein with small τ_e contributions are also highlighted in Figure 5.8a. In particular, Glu 197 and Glu 199 located in a 3_{10}-helix are also situated at the entrance of the active site.

Identification of the residues undergoing conformational exchange requires more careful scrutiny since the magnitude of the derived R_{ex} terms from the model-free analysis are rather small (Figure 5.7c). In this sense, the exchange terms were only taken into account when the ^{15}N relaxation data sets agreed at the three spectrometer frequencies (Figure 5.7c). By this criterion, only four regions of the protein are identified as undergoing motions in the μs-ms timescales. These include residues Phe 2

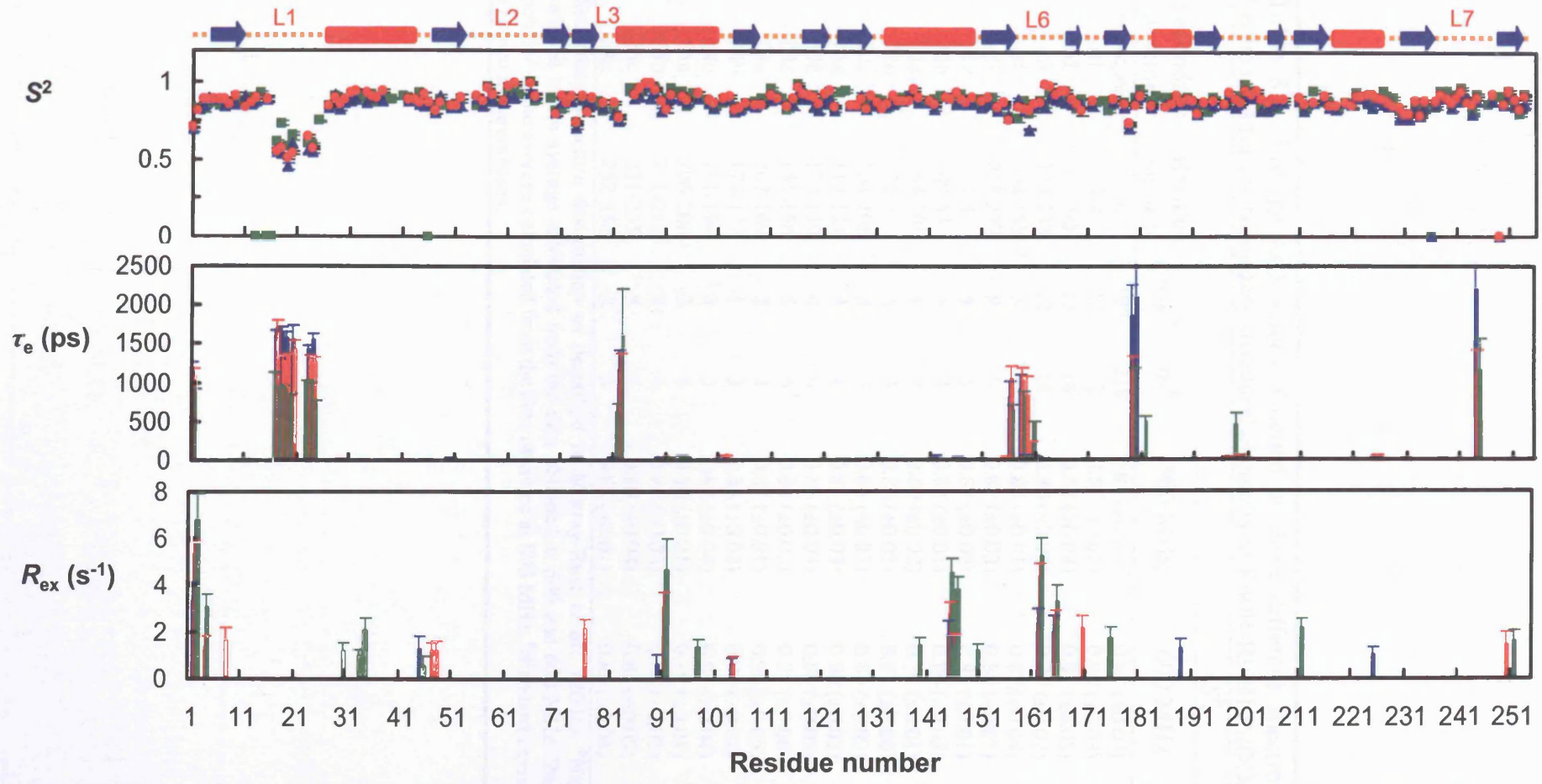

Figure 5.7. Model-free analysis of the internal motional parameters of R40E-R98H *Pa*DDAH. Values of the generalised order parameter S^2 (a), values of the fast internal motions τ_e (b) and values of the conformational exchange contributions R_{ex} (c) are shown as a function of residue number, extracted at three different magnetic field strengths. Plotted in blue are data extracted at ^{1}H nominal frequency of 500 MHz, in red at 600 MHz and in green at 800 MHz. Secondary structural elements are indicated at the top of the Figure. The designation L1 refers to the loop L1. Residues plotted on the abscissa in (a) could not be fit to any motional model.

Table 5.5. Average $\langle S^2 \rangle$ values obtained at three different spectrometer nominal frequencies for the secondary structural elements of R40R-R98H *Pa*DDAH

Secondary structure[a]	Residue number	$n_{5,6}$[b]	n_8[c]	$\langle S^2 \rangle$ 500 MHz	$\langle S^2 \rangle$ 600 MHz	$\langle S^2 \rangle$ 800 MHz
β/α propeller	All	195	219	0.85 (±0.07)[d]	0.87 (±0.07)	0.87 (±0.06)
α1	27-43	12	12	0.87 (±0.02)	0.90 (±0.03)	0.89 (±0.02)
α2	83-100	13	14	0.88 (±0.04)	0.91 (±0.05)	0.91 (±0.04)
α3	134-148	12	14	0.88 (±0.02)	0.90 (±0.02)	0.89 (±0.02)
α4	184-188	3	4	0.86 (±0.01)	0.87 (±0.04)	0.90 (±0.01)
α5	218-227	9	9	0.87 (±0.03)	0.88 (±0.03)	0.89 (±0.04)
β1a	4-8	5	5	0.87 (±0.02)	0.88 (±0.01)	0.85 (±0.01)
β1b	49-53	3	4	0.83 (±0.01)	0.86 (±0.03)	0.85 (±0.02)
β2a	68-70	1	2	0.80 (±0.02)	0.79 (±0.01)	0.85 (±0.06)
β2b	75-77	3	3	0.86 (±0.02)	0.87 (±0.06)	0.85 (±0.02)
β2c	104-106	3	3	0.83 (±0.03)	0.84 (±0.02)	0.84 (±0.03)
β3a	118-121	4	4	0.85 (±0.03)	0.90 (±0.03)	0.89 (±0.03)
β3b	124-129	4	6	0.86 (±0.03)	0.86 (±0.03)	0.86 (±0.04)
β3c	151-156	5	6	0.83 (±0.03)	0.85 (±0.06)	0.85 (±0.04)
β4a	167-169	3	3	0.87 (±0.01)	0.90 (±0.05)	0.88 (±0.06)
β4b	174-177	1	3	0.88 (±0.02)	0.90 (±0.01)	0.86 (±0.04)
β4c	191-194	3	3	0.84 (±0.04)	0.85 (±0.04)	0.84 (±0.04)
β5a	206-208	3	4	0.86 (±0.01)	0.89 (±0.01)	0.86 (±0.03)
β5b	211-215	3	4	0.86 (±0.03)	0.88 (±0.02)	0.88 (±0.05)
β5c	231-235	5	5	0.80 (±0.04)	0.83 (±0.05)	0.82 (±0.03)
β1c	252-254	3	3	0.85 (±0.01)	0.88 (±0.04)	0.86 (±0.06)

[a]Secondary structure designation as described in Murray-Rust *et al.*, (2001). [b]Number of residues included in the average calculated from the data obtained at 500 and 600 MHz. [c]Number of residues included in the average calculated from the data obtained at 800 MHz. [d]Standard deviation from average is shown in parentheses.

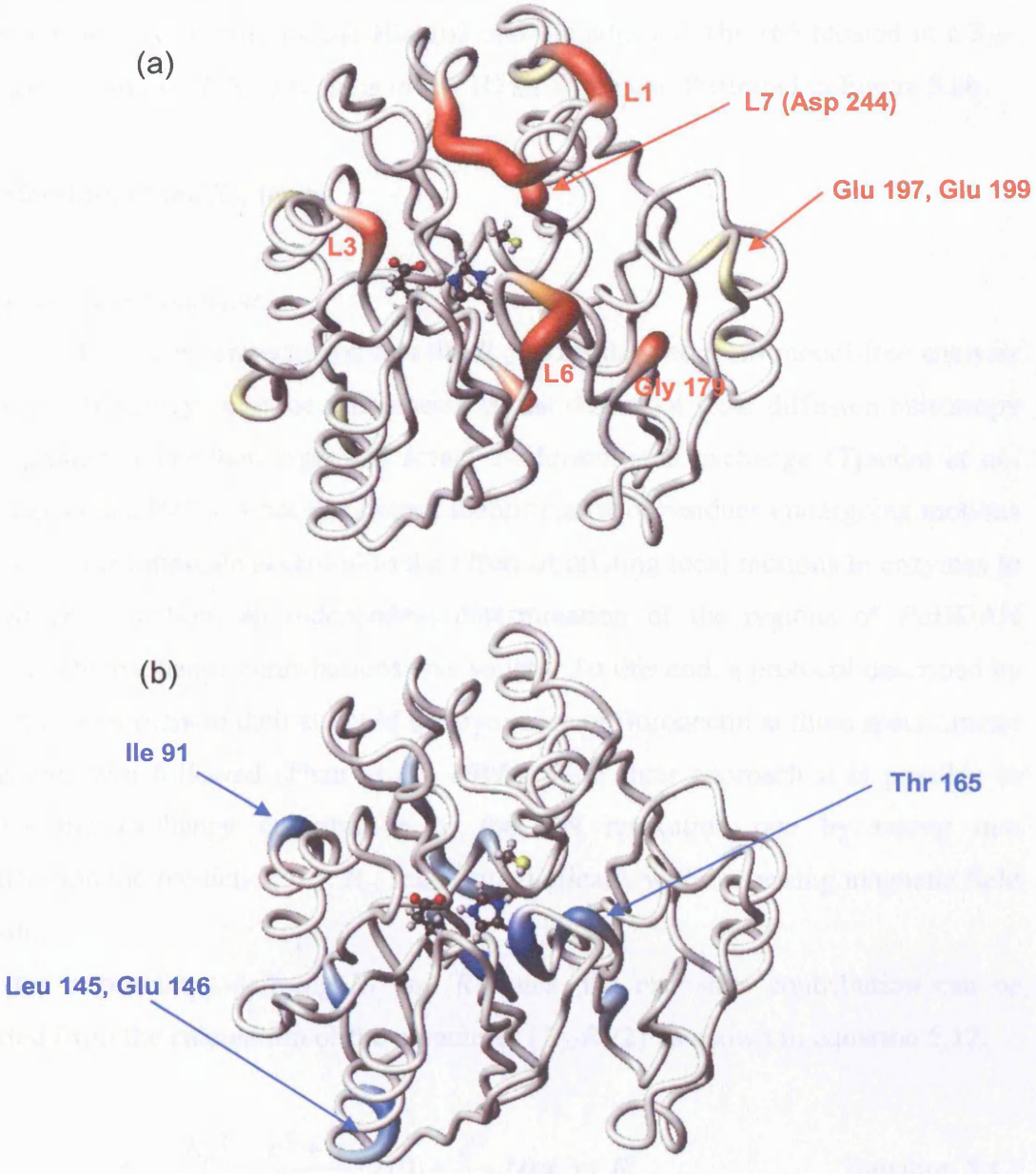

Figure 5.8. (a) Values of τ_e (ps) and (b) values of R_{ex} (s^{-1}) contributions estimated from the model-free analysis of the ^{15}N relaxation data acquired at 600 MHz have been plotted onto the X-ray structure of the monomeric *Pa*DDAH bound to L-citrulline (pdb code:1h70). The colour coding is as follows: in (a) coloured in red are residues for which $\tau_e > 1$ ns, in orange are residues for which 60 ps $> \tau_e > 1$ ns and in light yellow are residues for which $\tau_e < 60$ ps. In (b), coloured in dark blue are residues for which $R_{ex} > 5$ s^{-1}, in light blue are residues for which 1 $s^{-1} < R_{ex} < 5s^{-1}$. The active site residues Glu 114, His 162 and Cys 249 are shown in a ball-and-stick model. Note, residues 21 and 22 in L1 have not been assigned.

and His 4 located at the N-terminus, residue Ile 91 located at the bend of α-helix 2, residues Leu 145 and Glu 146 situated at the bottom part of α-helix 3, and more importantly the active site residue His 162 and the adjacent Thr 165 located in a 3_{10}-helix. The locations of these regions in the 3D structure are illustrated in Figure 5.8b.

5.4 Validation of the R_{ex} terms

Linear regression analysis

Several studies have demonstrated that the R_{ex} terms derived from model-free analysis at a single frequency could be reminiscent of the degree of local diffusion anisotropy of the protein rather than represent actual conformational exchange (Tjandra *et al.*, 1996; Lee *et al.*, 1997). Since the correct identification of residues undergoing motions in the μs to ms timescale is central to the effort of relating local motions in enzymes to the catalytic function, an independent determination of the regions of *Pa*DDAH influenced by exchange contributions was sought. To this end, a protocol described by Phan and co-workers in their study of the dynamics of fibronectin at three spectrometer frequencies was followed (Phan *et al.*, 1996). With their approach it is possible to quantify the exchange contribution to the ^{15}N relaxation rate by taking into consideration the prediction that R_{ex} scales quadratically with increasing magnetic field strength.

From the expressions defining R_1 and R_2 rates, the exchange contribution can be extracted from the calculation of the parameter $\{R_2\text{-}R_1/2\}$ as shown in equation 5.12.

$$\left\{R_2 - \frac{R_1}{2}\right\} = \frac{c^2 + d^2}{3} J(0) + \frac{d^2}{2} J(\omega_H) + R_{ex} \qquad \text{Equation 5.12}$$

where d and c have been defined previously in equations 5.6 and 5.7 respectively. Assuming equation 5.12 describes an exponentially decaying auto-correlation function and $R_{ex} > 1s^{-1}$, the second term of the equation which is field dependent is small relative to the others and can effectively be neglected. By introducing the term describing the quadratic dependence of $R_{ex} = AB_0^2$ into the equation, where A is a constant and B_0 the static magnetic field, equation 5.12 can be rearranged in the following manner.

$$\left\{R_2 - \frac{R_1}{2}\right\} \approx \left(\frac{c_1^2}{3}J(0) + A\right)B_0^2 + \frac{d^2}{3}J(0) \qquad \text{Equation 5.13}$$

where $c_1 = c/B_0$. When plotting $\{R_2\text{-}R_1/2\}$ as a function of B_0^2, $J(0)$ can be deduced from the intercept of the linear regression, *i.e.* $I_0 = d^2J(0)/3$ and the exchange constant A is given by the slope $m = [J(0)c_1^2/3] + A$:

$$A = m - \frac{c_1^2 I_0}{d^2} \qquad \text{Equation 5.14}$$

This equation predicts that any NH spin for which $m > (c_1^2 I_0/d^2)$ will have an exchange contribution and a value for R_{ex} can consequently be derived. However, due to experimental errors in data fitting and the limited number of data points in the linear regression analysis, A is not systematically equal to 0 for non-exchanging residues. In this sense, Phan *et al.* set a threshold value for the slope of $1.3\langle m\rangle$, where $\langle m\rangle$ is the average slope. The residues for which the regression analysis estimated a $m > 1.3\langle m\rangle$ were identified as subject to conformational exchange.

Using the treatment suggested by Phan *et al.* on the data obtained for *Pa*DDAH, 10 residues were identified as undergoing conformational exchange. The linear regression plots obtained for each of the 10 residues are shown in Figure 5.9. The values of R_{ex} obtained were compared to the values obtained from model-free analysis (Table 5.6).

Table 5.6. Comparison of chemical exchange rates derived from linear regression (LR) and model-free (MF) analyses

	R_{ex} (s^{-1}) at various B_0 (T)					
	11.74		14.09		18.79	
Residue	LR[a,b]	MF[a]	LR	MF	LR	MF
Phe 2	1.6 (±0.6)	3.4 (±0.7)	2.3 (±0.7)	5.4 (±0.5)	4.0 (±0.9)	6.8 (±1.1)
Ile 91	3.3 (±0.3)	N/A[c]	4.7 (±0.3)	3.1 (±0.6)	8.4 (±0.4)	4.6 (±1.3)
Gly 116	1.5 (±0.0)	N/A	2.1 (±0.0)	N/A	3.8 (±0.0)	N/A
Ile 142	2.1 (±0.2)	N/A	3.0 (±0.3)	N/A	5.4 (±0.3)	N/A
Leu 145	1.5 (±0.1)	1.8 (±0.7)	2.1 (±0.1)	2.8 (±0.5)	3.8 (±0.2)	4.5 (±0.6)
His 162	1.6 (±0.8)	2.3 (±0.6)	2.3 (±1.0)	4.4 (±0.5)	4.1 (±1.4)	5.2 (±0.8)
Lys 164	1.4 (±0.7)	N/A	2.0 (±0.8)	N/A	3.5 (±1.1)	N/A
Arg 241	1.6 (±0.5)	N/A	2.3 (±0.6)	N/A	4.1 (±0.7)	N/A
Ser 248	1.6 (±0.8)	N/A	2.3 (±0.9)	N/A	4.1 (±1.3)	N/A
Cys 249	3.9 (±1.2)	N/A	5.6 (±1.4)	N/A	9.9 (±1.8)	N/A
Ser 251	1.6 (±0.2)	N/A	2.4 (±0.3)	N/A	4.2 (±0.4)	1.6 (±0.5)

[a]LR stands for linear regression and MF for model-free. [b]$R_{ex} = AB_0^2$ in which A is defined in Equation 5.20.[c]Model-free analysis does not predict exchange contributions for the residue in question.

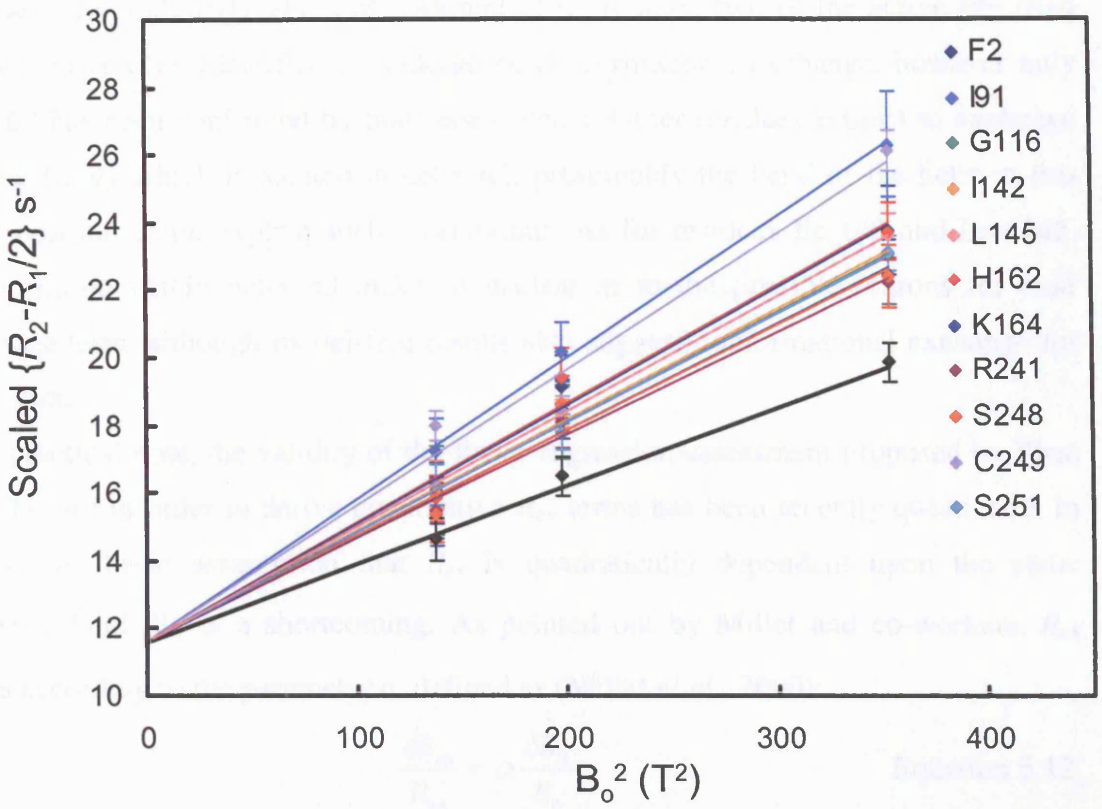

Figure 5.9. Linear regression curves of R_2-R_1/2 values plotted as a function of the static magnetic field B_0. Only the curves for which the slope m is >1.3*⟨m⟩ have been plotted and the corresponding residues are shown. The intercepts for all curves have been scaled to match the average value I_0 obtained for non-exchanging residues so that only the slopes of the linear regressions are compared. The regression curve for non-exchanging residues is shown in black.

For comparative reasons, the values of R_{ex} obtained from the linear regression and model-free analyses at 600 MHz are plotted onto the X-ray structure of *Pa*DDAH (Figure 5.10). Interestingly, the majority of residues undergoing conformational exchange are located around the active site region in short loops or 3_{10}-helices, such as residues 116, 162, 164, 241, 248, 249 and 251. Notably, two of the active site triad residues have been identified as undergoing conformational exchange, however only His 162 has been confirmed by both assessments. Other residues subject to exchange include Ile 91 which is located in helix α2, presumably the bend of the helix at this exact position could explain such a behaviour. As for residues Ile 142 and Leu 145, their position within helix α3 makes it unclear as to the possible reasons for their exchange term, although model-free results also suggest conformational exchange for this region.

On a practical note, the validity of the linear regression assessment proposed by Phan *et al.* (1996) in order to derive quantitative R_{ex} terms has been recently questioned. In reality, the basic assumption that R_{ex} is quadratically dependent upon the static magnetic field B_0 is a shortcoming. As pointed out by Millet and co-workers, R_{ex} scales according to the parameter α, defined as (Millet *et al.*, 2000):

$$\frac{\delta R_{ex}}{R_{ex}} = \alpha \frac{\delta B_0}{B_0} \qquad \text{Equation 5.12}$$

$$\alpha = \frac{\mathrm{d} \ln R_{ex}}{\mathrm{d} \ln \Delta\omega} \qquad \text{Equation 5.13}$$

Where $\Delta\omega$ is the chemical shift difference between exchange sites A and B and α satisfies the constraints $0 \leq \alpha \leq 2$, providing the population of site A is much larger than the population of site B ($p_a >> p_b$). In which case, α takes the following form:

$$\alpha = \frac{2(k_{ex} / \Delta\omega)^2}{1 + (k_{ex} / \Delta\omega)^2} \qquad \text{Equation 5.14}$$

where k_{ex} is the exchange rate constant between sites A and B. From these equations R_{ex} will only scale quadratically with the applied static field if α = 2, which would indicate that all exchange contributions are in the time fast exchange limit relative the chemical shift time scale. In reality this result has been proven not to be the case for most NH spins in basic bovine pancreatic trypsin inhibitor (BPTI) when examined by ^{15}N CPMG measurements of transverse relaxation (Millet *et al.*, 2000). Despite this caveat, the justification for assuming that R_{ex} scales quadratically with B_0 is merely

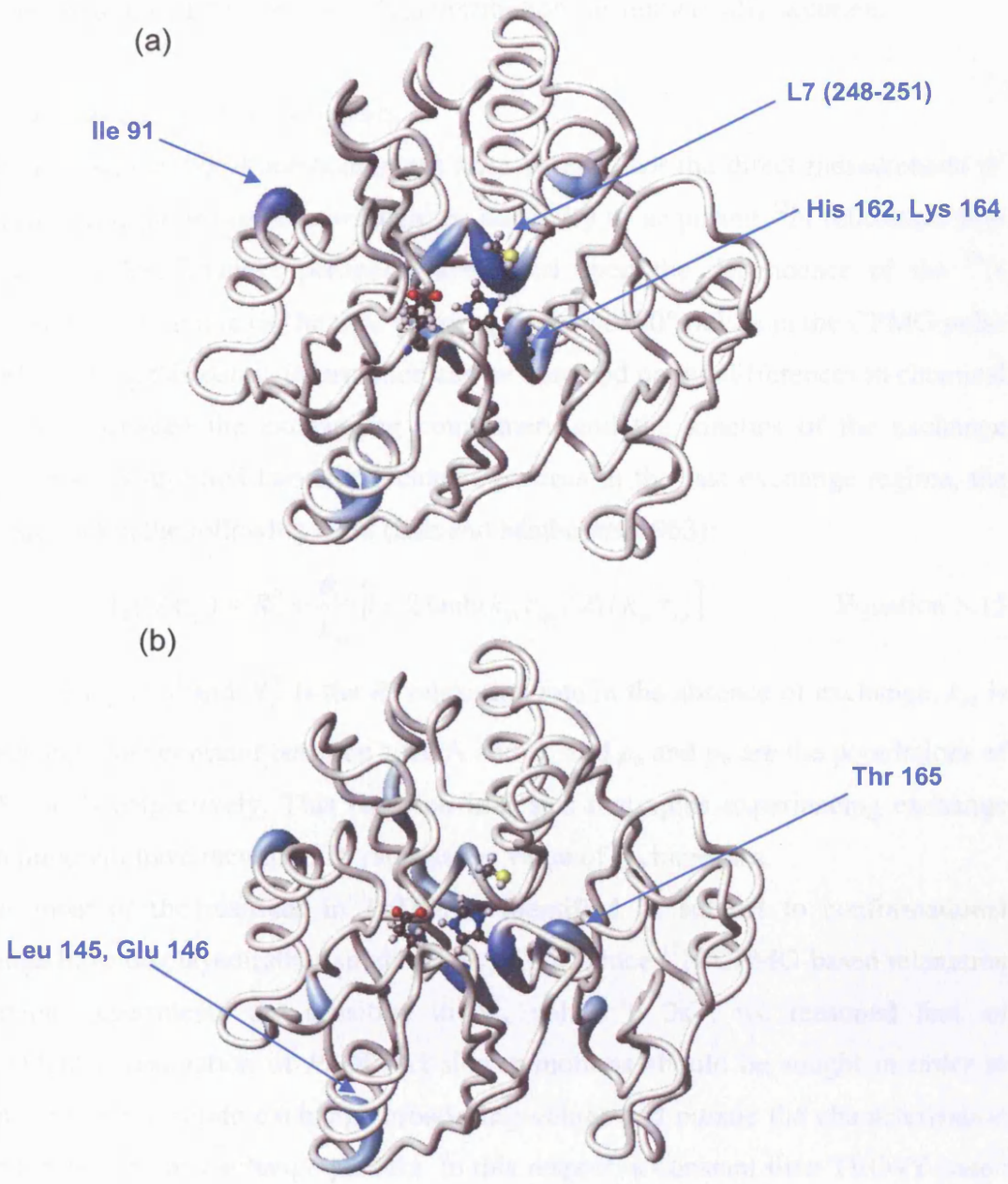

Figure 5.10. Values of R_{ex} (s^{-1}) contributions estimated from the linear regression analysis (a) and model-free analysis (b) of the ^{15}N relaxation data acquired at 600 MHz have been plotted onto the X-ray structure of the monomeric PaDDAH (pdb code:1h70). The colour coding is as follows: coloured in dark blue are residues for which.$R_{ex} > 5\ s^{-1}$, in light blue are residues for which $R_{ex} < 5s^{-1}$. The width of the tube is also proportional to the R_{ex} contributions. The active site residues Glu 114, His 162 and Cys 249 are shown in a ball-and-stick model.

qualitative at this stage. Since ^{15}N relaxation is available at three magnetic field strengths our purpose was mainly to identify possible residues undergoing exchange, however not to regard the values of R_{ex} contributions as numerically accurate.

^{15}N relaxation dispersion experiments

Recent advances in NMR methodologies have allowed for the direct measurement of chemical-exchange phenomena without the necessity of acquiring ^{15}N relaxation data at multiple fields. These experiments are based upon the dependence of the ^{15}N transverse relaxation rate on the time delay τ_{cp} between 180° pulses in the CPMG pulse train. By varying this delay, information can be obtained on the differences in chemical shifts ($\Delta\omega$) between the exchanging conformers and the kinetics of the exchange process. For a simplified two-site exchange process in the fast exchange regime, the expression takes the following form (Luz and Meiboom, 1963):

$$R_2(1/\tau_{cp}) = R_2^0 + \frac{\phi_{ex}}{k_{ex}}\left[1 - 2\tanh(k_{ex}\tau_{cp}/2)/k_{ex}\tau_{cp}\right] \qquad \text{Equation 5.15}$$

where $\phi_{ex} = p_a p_b \Delta\omega^2$ and R_2^0 is the R_2 relaxation rate in the absence of exchange, k_{ex} is the exchange rate constant between sites A and B, and p_a and p_b are the populations of site A and B respectively. This equation indicates that spins experiencing exchange broadening will have increased R_2 rates as the value of τ_{cp} increases.

So far most of the residues in *Pa*DDAH identified as subject to conformational exchange have displayed rather small values of R_{ex}. Since ^{15}N CPMG-based relaxation dispersion experiments are sensitive to R_{ex} values > $3s^{-1}$, we reasoned that an independent investigation of *Pa*DDAH slower motions should be sought in order to determine more accurate exchange broadening values and pursue the characterisation of the kinetics of the exchange process. In this respect, a constant time TROSY-based CPMG ^{15}N relaxation dispersion experiment was acquired at a spectrometer frequency of 800 MHz (see Chapter 2, Section 2.4; Loria *et al.*, 1999; Tollinger *et al.*, 2001). Effective transverse relaxation rates (R_2^{eff}) were determined from the relation shown in Equation 5.16 (Mulder *et al.*, 2001):

$$R_2^{eff} = -\ln\left(\frac{I(\nu_{cpmg})}{I(0)}\right)/T \qquad \text{Equation 5.16}$$

where $I(\nu_{cmpg})$ represents the cross peak intensity at a particular CPMG pulsing frequency. In this case $\nu_{cmpg} = 1/\tau_{cp}$ was varied from 100 to 1000 s^{-1}. $I(0)$ is the peak

intensity obtained without the CPMG pulse train. T is the total relaxation time, in this case 40 ms. In order to identify *Pa*DDAH residues displaying R_2 rates with a significant amount of relaxation dispersion, the plot of $\{R_2^{eff}(1/\tau_{cp} = 100\ s^{-1}) - R_2^{eff}(1/\tau_{cp} = 1000\ s^{-1})\}$ as a function of residue number is shown in Figure 5.11. The first step in interpreting the results is a coarse filtering of the data. From experience, residues that display R_2 rates which do not differ by more than 3 s^{-1} over the entire 900 s^{-1} range of $1/\tau_{cp}$ values are generally not affected by conformational exchange. The filtering procedure is necessary since the noise level of the experiment is close to $\pm$ 3 s^{-1}, which is consistent with what is observed in other studies of this kind (Beach *et al.*, 2005; Valentine and Palmer, 2005; Wang *et al.*, 2004).

From the inspection of ΔR_2^{eff} values, no particular residue of *Pa*DDAH displays a significant amount of exchange contribution to the R_2 rate. In fact, residues previously identified through model-free and linear regression analyses as exhibiting exchange terms of 2 to 3 s^{-1} do not stand out in this experiment. Only one residue Gln 140 has a value of $\Delta R_2^{eff} > 3\ s^{-1}$, however no attempt has been made to fit the decay curve of $R_2(1/\tau_{cp})$, as this experiment is a first effort in our laboratory to implement ^{15}N relaxation dispersion experiments and the confirmation of this result is still the subject of on-going research.

5.5 Reduced spectral density function analysis

An alternative approach to the analysis of ^{15}N relaxation measurements is the reduced spectral density procedure initially described by Farrow *et al.* (1995). The advantage of this approach over the model-free formalism is that the information on the timescales of protein motions can be extracted without the need to assume a model for the overall tumbling of the molecule or the exact form of the spectral density function. Instead, it is possible to estimate $J(\omega)$ at seven different frequencies when ^{15}N relaxation data is available at three different magnetic field strengths. In this case, $J(\omega_N)$ and $J(0.87\omega_H)$ estimated at 500, 600 and 800 MHz proton frequencies and $J(0)$, were calculated from the relaxation data obtained for the monomeric *Pa*DDAH from the following equations:

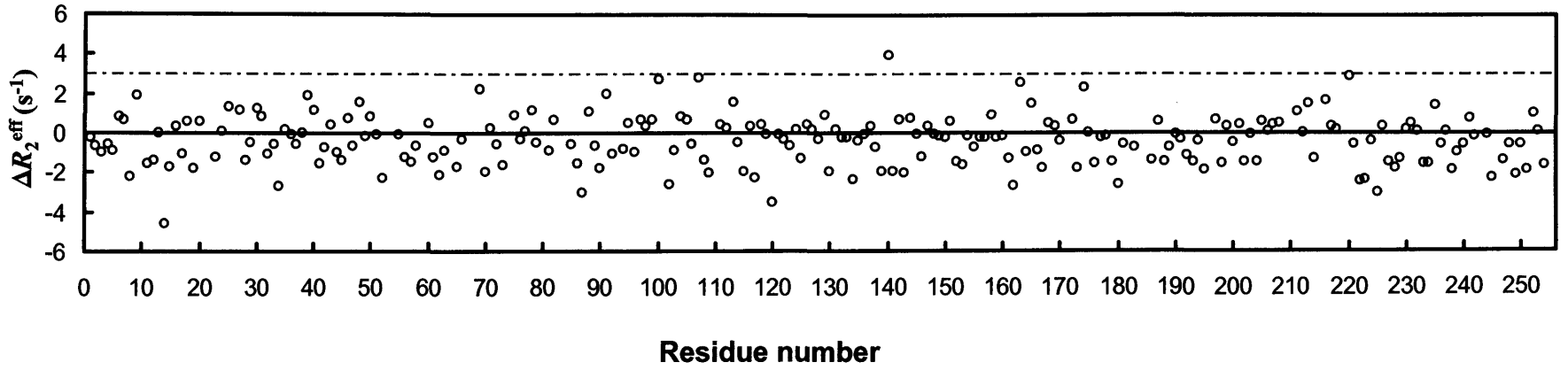

Figure 5.11. ΔR_2^{eff} values plotted as a function of residue number. The dashed line indicates the 3 s^{-1} threshold line (see text).

$$J(0) = \left[R_2 - (3d^2/4 + c^2)J(\omega_N) - (13d^2/8)J(0.87\omega_H)\right]/(d^2 + 4c^2/3) \quad \text{Equation 5.17}$$

$$J(\omega_N) = \left[R_1 - (7d^2/4)J(0.87\omega_H)\right]/\left[3d^2/4) + c^2\right] \quad \text{Equation 5.18}$$

$$J(0.87\,\omega_H) = (4/5d^2)(\gamma_N/\gamma_H)(\text{NOE} - 1)R_1 \quad \text{Equation 5.19}$$

where d and c have been defined in Equations 5.4 and 5.5 respectively. The derived reduced spectral densities for *Pa*DDAH are plotted on a per-residue basis in Figure 5.12.

The $J(0)$ term in the reduced spectral density analysis is known to report on motions in the slower motional timescales (μs-ms) and is hence perturbed by chemical exchange contributions present in the transverse relaxation rates (Farrow *et al.*, 1995). In order to correctly estimate the value of $J(0)$ free from chemical exchange contributions, the relaxation data acquired at multiple field strengths can be used to quantify the R_{ex} term in a per-residue-basis manner. However this procedure has not been undertaken here, since it is based on the previous assumption that R_{ex} should scale quadratically with B_0 (Farrow *et al.*, 1995). More recently, other methods to calculate $J(0)$ have been put forward by Kroenke and co-workers in which the longitudinal cross-correlation relaxation rate constants η_z and η_{xy} for ^1H-^{15}N dipolar and ^{15}N CSA relaxation interference are measured (Kroenke *et al.*, 1998). In contrast to R_2/R_1 ratios, η_{xy}/η_z ratios are independent of chemical exchange processes and report solely on internal and overall motions that contribute to relaxation. Ultimately, $J(0)$ can then be expressed as:

$$J(0) = \frac{3}{4}\left[2\frac{\eta_{xy}}{\eta_z} - 1\right]J(\omega_N) \quad \text{Equation 5.20}$$

However no attempt was made to determine cross-correlation rates and the $J(0)$ values reported here at 600 and 800 MHz for individual NH spins have not been corrected for chemical exchange (Figure 5.12a). $J(0)$ values were directly derived from Equation 5.17. From the inspection of Figure 5.12a, both data set values are almost identical with the exception of very few residues. For our purposes, values of $J(0)$ one standard deviation higher than the average have been interpreted to indicate the presence of motions operating on the microsecond and millisecond timescales. In the case of *Pa*DDAH, from the inspection of $J(0)$ values, residues Phe 2, Ile 91, Gly 116 and the regions from residues 61-65, 145-146 and residues 162-165 have been identified as

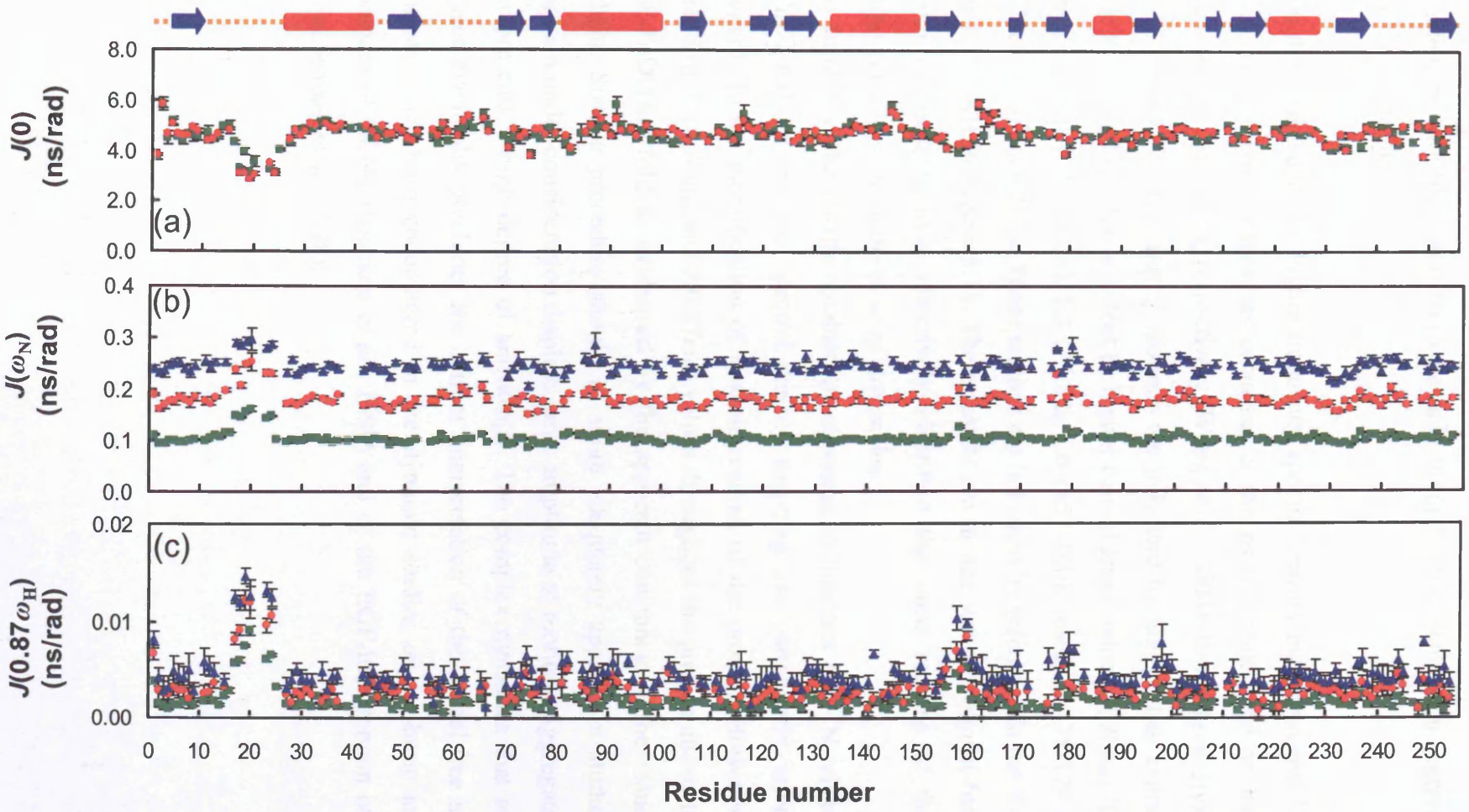

Figure 5.12. Values of the reduced spectral densities $J(\omega)$ of R40E-R98H *Pa*DDAH estimated at (a) $J(0)$, (b) $J(\omega_N)$ and (c) $J(0.87\omega_H)$ plotted as a function of residue number and extracted at three different magnetic field strengths (blue triangle at ^{1}H nominal frequency of 500 MHz, red circles at 600 MHz and green squares at 800 MHz). Secondary structural elements are indicated at the top of the Figure. $J(0)$ was not estimated with data acquired at 500 MHz.

subject to slow motions. Apart from residues 61-65 which lie in the L2 loop inside the active site cavity, the other regions presenting motions in the slower timescales are in agreement with the results from the field strength dependent linear regression analysis presented earlier.

On the other hand, the higher frequency spectral density measurements i.e. $J(0.87\omega_H)$ and $J(\omega_N)$ report on motions occurring at the ps timescales and on motions at the nanosecond timescales, respectively (Wang *et al.*, 2004). From the analysis of Figures 5.12b and 5.12c, the 'faster' motions are indicated by higher than average values of $J(0.87\omega_H)$ and to a lesser extent by higher than average values of $J(\omega_N)$. These regions include loops L1 (13-24), L3 (79-82), L6 (157-160), residues 178-179 and 197-199 which lie in 3_{10}-helices. These regions are indicated by reference to the X-ray structure of *Pa*DDAH in Figure 5.13. These results are in line with the model-free analysis as both assessments have effectively identified the same regions of the protein as inherently flexible in the ps to ns timescales.

In summary, the results obtained by choosing to interpret the ^{15}N relaxation data of *Pa*DDAH by reduced spectral density mapping are particularly informative with regards to the identification of flexible regions of the protein. However, the overall trend of $J(0)$, $J(\omega_N)$ and $J(0.87\omega_H)$ values throughout the protein attests to the rigidity of *Pa*DDAH fold as evidenced by the apparent constancy of the values across the chain. Such a procedure should be more adequately applied in studies where the protein under consideration displays large amplitude of motions, aggregates in solution or presents a high degree of anisotropy. The complex dynamics that arises in such cases eventually precludes the correct interpretation of the model-free analysis. This situation has been encountered in the dynamic studies of the basic leucine zipper domain of GCN4 (Bracken *et al.*, 1999) and of the EGF-like domain of heregulin-α (Fairbrother *et al.*, 1998).

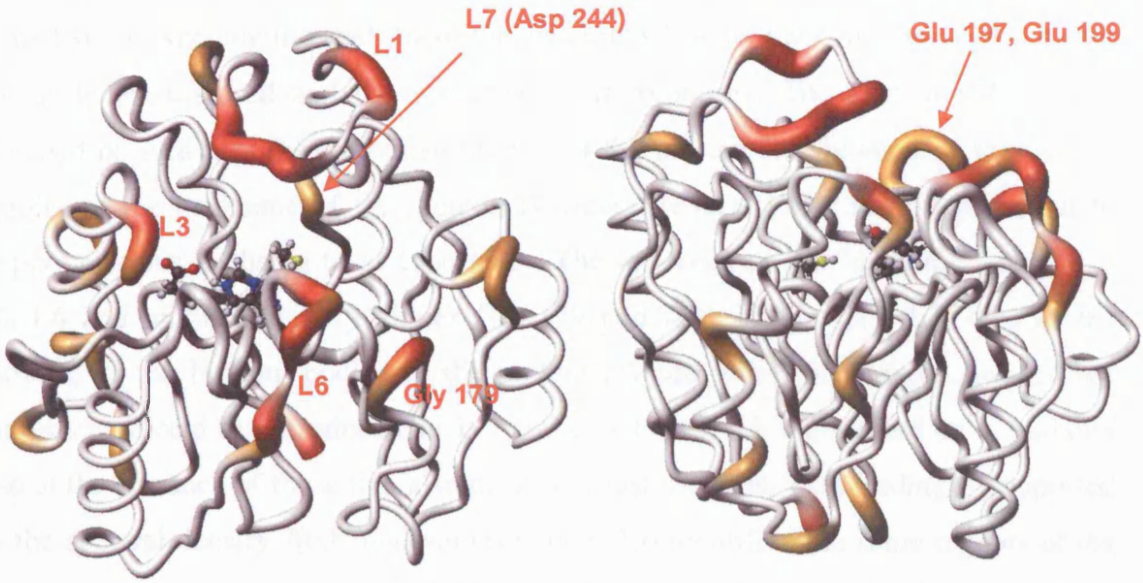

Figure 5.13. Front and lateral views of the 3D structure of *Pa*DDAH showing in red residues for which the estimated values of $J(0.87\omega_H)$ are three standard deviations (SD) higher than the average value of $J(0.87\omega_H)$ estimated at 600 MHz. Residues for which $J(0.87\omega_H)$ > 1*SD are shown in orange. The width of the tube is also proportional to the $J(0.87\omega_H)$ values. The active site residues Glu 114, His 162 and Cys 249 are shown in a ball-and-stick model. Note, residues 21 and 22 in L1 have not been assigned.

5.6 Discussion

The backbone dynamics of the engineered monomeric *Pa*DDAH enzyme have been investigated through the measurement of ^{15}N R_1, R_2 and $\{^1H\}$-^{15}N heteronuclear relaxation parameters. The acquisition of data at multiple magnetic field strengths has enabled a more reliable identification of the regions of the protein susceptible to undergo motions at multiple timescales.

The three data sets obtained were analysed independently through the model-free formalism. In keeping the analysis of the data sets separate, the consistency of the data sets is better assessed and the χ^2 statistics are kept low. The average $\langle S^2 \rangle$ value obtained of ~0.87 for the structured regions of the protein clearly attest to the overall rigidity of the backbone of the protein. However, several loops have been found to display motions in the ps to ns timescales. The locations of the "dynamic" loops L1, L3, L6 and of the isolated residues Gly 179 and Asp 244 in loop L7, map to the entrance of the binding pocket of the protein (Figure 5.8a). In addition, model-free analysis predicted that residues Glu 197 and Glu 199 which are located on a 3_{10}-helix also at the entrance of the active also displayed fast motions. This finding is supported by the spectral density $J(0.87\omega_H)$ values which also identified the same regions of the protein as undergoing dynamics in the faster timescales (Figure 5.13). Taken together, these results suggest that the L1 loop which is subject to the largest amplitude conformational fluctuations in the protein may play an important role in facilitating substrate intake and product release. Similarly, the sampling of fast multiple alternate states for loops at the entrance of the binding cavity could dictate an easier access and release of the ligand. This mechanism has been seen for other enzymes such as in the metallo-β-lactamase for which the fast dynamics of the major flap and the minor loop at the entrance of the active site are abolished upon inhibitor binding, alluding to a locking mechanism around the substrate (Huntley *et al.*, 2000). Another example comes from the extended studies of the enzyme dihydrofolate reductase and the importance of the different motional timescales sampled by the "Met-loop" over the active site during binding, catalysis and release of the ligands (Osborne *et al.*, 2001; McElheny *et al.*, 2005). The Met-loop also samples ps to ns timescales in the apo-state of the enzyme. More interestingly, the studies of the HIV protease by Torchia and co-workers have shown that the flaps that form part of the substrate binding site sample

sub-ns fluctuations in the free protease (Freedberg *et al.*, 2002). The authors of this work have inferred that these motions reflect in reality a dynamic equilibrium between semi-open conformations of the protein. This surmise could be applied to *Pa*DDAH, in which case the fast sampling of conformations of the L1 loop over the active site, would corroborate crystallographic data. No defined electron density was found for residues 14-25 for the apo-form of the protein (Murray-Rust *et al.*, 2001). Since the presence of loops over active site regions is not unprecedented, we can speculate about their functional role as able to permit an induced fit between substrate and enzyme.

The slower motional timescales have also been investigated for *Pa*DDAH. Since several enzymatic reactions occur in the microsecond to millisecond timescales, it has been inferred that protein motions at these timescales are pivotal to catalysis (Fersht, 1999). For this reason, particular attention was paid to the optimisation of rotational diffusion model of the enzyme, as protein conformational exchange determined through the model-free analysis may be influenced by the underestimation of the anisotropy of the rotational tumbling of the molecule. For comparative purposes, ^{15}N relaxation data obtained at 500 MHz was additionally analysed assuming isotropic tumbling of the protein. The analysis of the residues which presented R_{ex} terms showed that helices α1, α3, and α5 exhibited significant conformational exchange (data not shown). These exchange contributions are absent when the rotational diffusion of *Pa*DDAH is taken, more correctly, to be axially symmetric. Overall only a very small number of residues could be confidently identified as undergoing exchange contributions from the model-free results (Figure 5.8a). These included the active site residue His 162 and residues around the active site such as Thr 165. Building on these findings, additional protocols were sought in order to confirm their veracity. Although qualitative, the linear regression analysis performed of $\{R_2\text{-}R_1/2\}$ values as a function of B_0^2 together with the spectral density analysis of $J(0)$ gave a more complete view of the regions potentially undergoing slower motions (Figure 5.10). Interestingly, these included four regions around the active site, with Gly 116 lying in the same 3_{10}-helix as the active site residue Glu 114, residues 61-65 which form the L2 loop, residues 162-165 located adjacent to the L6 loop and the region that comprises the third active site residue, the L7 loop (248-251). The analysis of the protein-ligand contacts based on the X-ray structure of *Pa*DDAH complexed with L-citrulline has shown that all

these residues make non-bonded interactions with the ligand. This view is supported by independent studies in which R_{ex} terms have been found around the active site region of other enzymes, in particular those of free ribonuclease binase and free RNase (Wang *et al.*, 2001; Beach *et al.*, 2005). In these studies the μs to ms motions were preserved both in the apo- and substrate mimicked state of the enzyme. The authors allude to the notion that the free enzymes sample both apo- and bound conformations and that the binding of the ligand simply shifts the equilibrium towards the bound conformer. A detailed analysis of *Pa*DDAH in the bound state is necessary in order to assess the effects of the ligand on the backbone dynamics and whether the slow dynamics are also preserved in both states.

From a kinetic perspective, enzymes are known to undergo dynamic processes implicit in their catalytic cycle. The various investigations of enzyme dynamics have been able to link the rates of protein conformational rearrangements occurring both in the apo- and bound forms of the enzyme to the rates of catalytic turnover. The limiting step in the catalytic process is in fact protein movement. With this in mind, we also sought to gain insight into the kinetics of the conformation exchange process in the monomeric *Pa*DDAH. The use of ^{15}N relaxation dispersion experiments would have ideally clarified the significance of many of the small R_{ex} terms observed in the apo-state of *Pa*DDAH. So far our results are inconclusive and further attempts are needed to determine the best choice of pulse sequence. Ultimately, it is our understanding that - ^{15}N relaxation dispersion experiments are in fact crucial to the full understanding of *Pa*DDAH intrinsic dynamics and its influence upon the catalytic rates. In this vein the dynamics of the ligand-bound state of the enzyme still need to be addressed.

In the next chapter, the bound-state of the monomeric *Pa*DDAH enzyme will be investigated. Assessing the effects of ligand addition on *Pa*DDAH loop structure and mobility will give additional valuable structural insights. In collaboration with the clinical pharmacology department at UCL, a library of small chemical compounds known to inhibit both the human DDAH I and bacterial DDAH enzymes have been used to assess different binding schemes. Due to the overall weak protein-ligand interactions as revealed by the HSQC-based titration experiments and chemical shift mapping of the binding site, molecules able to react with the active site cysteine residue and remain covalently linked in the active site cavity were also investigated.

Chapter 6

Ligand binding studies of the monomeric *Pa*DDAH

Abstract

This chapter describes the use of small molecule ligands in the characterisation of the bound-state of the bacterial *Pa*DDAH enzyme. These compounds have been derived as part of a collaborative effort between several departments at UCL which sought to derive specific inhibitors against mammalian DDAH homologues. NMR methodology was sought as a means to map the binding site of the bacterial protein through chemical shift perturbation. Due it the sensitivity to weak interactions, heteronuclear NMR titration experiments were particularly useful to estimate dissociation constants for both substrate and product complexes. A general observation obtained from these titrations is that ligands that act as inhibitors of the enzyme *in vitro* give rise to specific broadening of a subset of N-H cross peaks corresponding to residues around the active site region and in the L1 loop that closes the ligand binding pocket. The line broadening is also observed with a ligand covalently bound to Cys249 of the catalytic triad. These results are interpreted to imply that in the bound state the dynamic profile of the backbone of the *Pa*DDAH enzyme is altered, probably with ordering of the L1 loop but retaining a component of conformational exchange that gives rise to residual line broadening, even in the saturated state. These findings challenge the current view of the L1 loop locked into position over the active site upon ligand binding which was given by the X-ray structure.

6.1 NMR spectroscopy applied to the study of protein-ligand interactions

The study of protein-ligand interactions by NMR is based upon differences in physico-chemical properties existing between the free and bound states of the protein under consideration. Commonly, such differences are interpreted in terms of an overall increase of the relative mass between the two states, which in turn cause changes in the diffusion and tumbling properties of the molecule. When monitoring the ligand, these changes can effectively be followed by specific NMR experiments which exploit the differential relaxation rates of the free ligand relative to the complexed form and rely on the transfer of magnetisation from the ligand to the protein and vice-versa (Pellecchia, 2005). The advantages of monitoring the ligand are the small quantity of target unlabelled protein necessary to perform the experiments, together with the possibility of screening multiple ligands in the same experiment. Alternatively, NMR spectroscopy can also monitor specific spin systems of the protein, which are sensitive to the chemical environment in which they lie. Perturbation of this environment, *e.g.* by the addition of a small molecule ligand in direct contact with the spin, translates into deviations in the chemical shifts of the spin observed in a particular NMR experiment. Providing that the resonance assignments for the protein are available, the observation of changes in the ^{1}H-^{15}N, ^{1}H-^{13}C HSQC- or TROSY-based correlation spectra upon ligand addition has become an important method for the determination of binding sites in proteins. Chemical shift mapping can in addition provide an estimate of the dissociation constants (K_d) of the protein-ligand complexes. When screening for lead compounds, this approach is particularly useful to detect low or high affinity binding, which has led to the development of NMR-based strategies for drug discovery, such as the structure-activity relationship by NMR (SAR by NMR) (Shuker *et al.*, 1996).

In line with these applications, the characterisation of a small library of compounds designed to inhibit the human enzyme DDAH I *in vitro* was sought. Since the mammalian DDAH enzymes are not obtained in suitable amounts for NMR characterisation, the screening of these inhibitors for potential binding affinity towards the bacterial enzyme seemed reasonable. In fact, *Pa*DDAH was recently suggested to be a potential antibacterial target (Vallance *et al.*, 2005), which validates the need to further assess the different binding specificities these compounds exhibit towards the DDAH enzyme family. However, for completeness the binding site of the monomeric

*Pa*DDAH will be investigated at first through titration experiments of the substrate N^{η},N^{η}-dimethylarginine (ADMA) and the product L-citrulline.

6.2 Studying the bound state of the monomeric *Pa*DDAH

Determination of the Michaelis-Menten constant K_M of the monomeric PaDDAH enzyme

In the course of our investigations of the structure-function characteristics of the DDAH enzyme, we sought to confirm that the two point mutations R40E and R98H in the *Pa*DDAH interface had not significantly altered the affinity and kinetic parameters of the monomeric enzyme relative to the WT protein. Recent studies on unrelated enzymes have revealed that mutations of distal residues could still have effects on enzymatic catalysis without interacting directly with the substrate (Rajagopalan *et al.*, 2002; Kovrigin and Loria, 2006), which is why a confirmation of catalytic rates of the *Pa*DDAH monomer was pursued.

Following the protocol presented in Chapter 2, Section 2.4, the K_M value of R40E-R98H *Pa*DDAH for the substrate N^{η},N^{η}-dimethylarginine (ADMA) has been determined to be ~460 μM. This value was compared to the published K_M value of WT *Pa*DDAH for ADMA which is in the order of 310 μM (Stone *et al.*, 2005a). The difference observed could attest to an overall decrease in affinity of the monomeric enzyme for ADMA, however the conditions used in the assays differ in terms of temperature and salt concentrations which could account for the difference observed. Nevertheless the k_{cat} values for both enzymes remain comparable, Stone *et al.* determined a k_{cat} value of 0.56 s^{-1} for WT *Pa*DDAH and the value obtained for the monomeric enzyme under the conditions described in Chapter 2, Section 2.4 is ~ 0.42 s^{-1}. It has to be noted that the catalytic turnover of *Pa*DDAH enzymes is extremely slow when compared to other well studied enzymes such as RNase A or ribonuclease binase. These enzymes have k_{cat} values in the range of 1400-2900 s^{-1} for RNase (del Cardayre *et al.*, 1995) and 1500 s^{-1} for ribonuclease (Schulga *et al.*, 1998).

Titration studies of ADMA, L-citrulline

In order to assess the ligand-bound state of *Pa*DDAH and effectively map the binding site of the enzyme, the molecules initially chosen for NMR titration experiments were the substrate ADMA and the product L-citrulline. The transition from the free to the ligand-bound state of the enzyme was followed through the acquisition of a series of [^{1}H-^{15}N]-HSQC spectra recorded as a function of increasing ligand concentration. Details of the experiment conditions can be found in Table 2.4 (Chapter 2, Section 2.5). Selected regions of the [^{1}H-^{15}N]-HSQC spectra acquired for each titration point illustrate the similar pattern of chemical shift perturbation obtained upon the binding of each ligand (Figures 6.1 and 6.2). In reality, the titration of the substrate into the active enzyme results in the characterisation of the binding site of the hydrolysis product L-citrulline. The catalytic reaction is bound to be completed prior to the acquisition of the HSQC experiment. Ultimately, titrating L-citrulline into the protein served as a control experiment to confirm that the chemical shift changes obtained upon the binding of ADMA are in fact the outcome of the presence of L-citrulline in the active site. In fact, L-citrulline is known to inhibit *Pa*DDAH activity at concentrations higher than 3 mM which would indicate that it can effectively bind to the active site (Stone *et al.*, 2005a). Moreover, it was additionally shown by the X-ray structure of the WT protein that the enzyme does crystallise in complex with the product (Murray-Rust *et al.*, 2001).

The chemical shift deviations (Δδ) upon ligand titration have been calculated from the following equation $\Delta\delta = \sqrt{(\Delta\delta_{^{15}N}/5)^2 + (\Delta\delta_{^{1}H^N})^2}$ and have been plotted as a function of residue number in Figure 6.3a. Initial inspection of these results clearly shows that the pattern of changes in chemical shifts upon the titration of both molecules is essentially identical except for the slightly larger chemical shift deviations observed during the titration of the substrate.

In addition to the chemical shift perturbations, it is important to note that for a subset of $^{1}H^{N}$-^{15}N cross peaks, the chemical shift changes could not be followed during the entire titration and no end point could be assigned. The ligand-dependent line broadening of these resonances is extreme and the intensity of the cross peaks decreases to noise levels early on in the titration which subsequently impairs the characterisation of the chemical shift perturbation. These residues are shown in Figure 6.3a in yellow bars. On the other hand several residue titration end points could not be

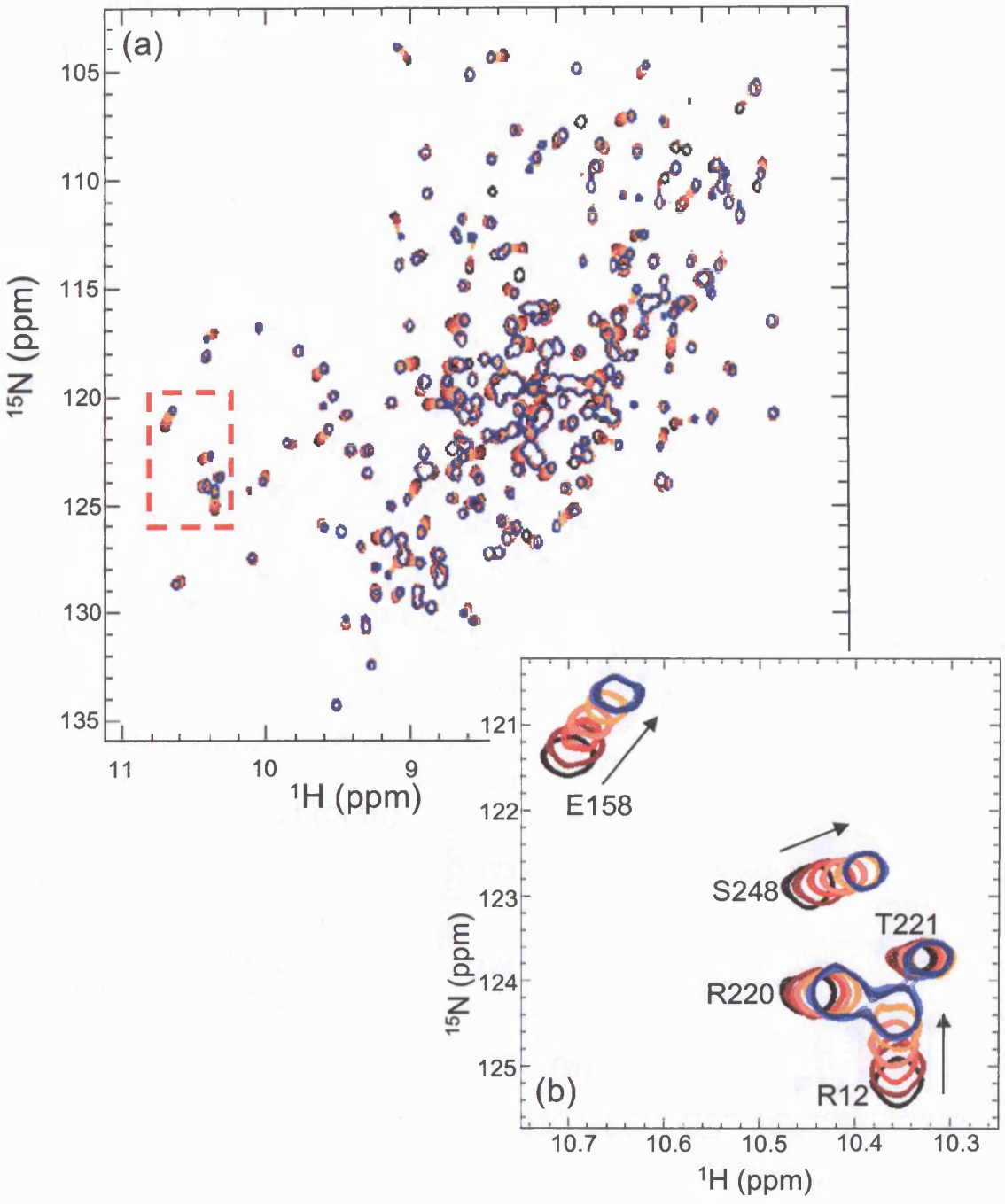

Figure 6.1. (a) Superimposed [^{1}H,^{15}N]-HSQC spectra of the monomeric [^{1}H,^{15}N]-labelled *Pa*DDAH acquired after the incremental addition of ADMA. The black spectrum corresponds to the free protein and the dark blue spectra corresponds to the last titration point upon saturation of the enzyme. (b) Inset of (a) as delimited by the red dotted box. The arrows indicate spectra acquired at increasing ADMA concentrations. The spectra were acquired at ^{1}H proton frequency of 500 MHz at 25°C (see Chapter 2, Section 2.5 for details).

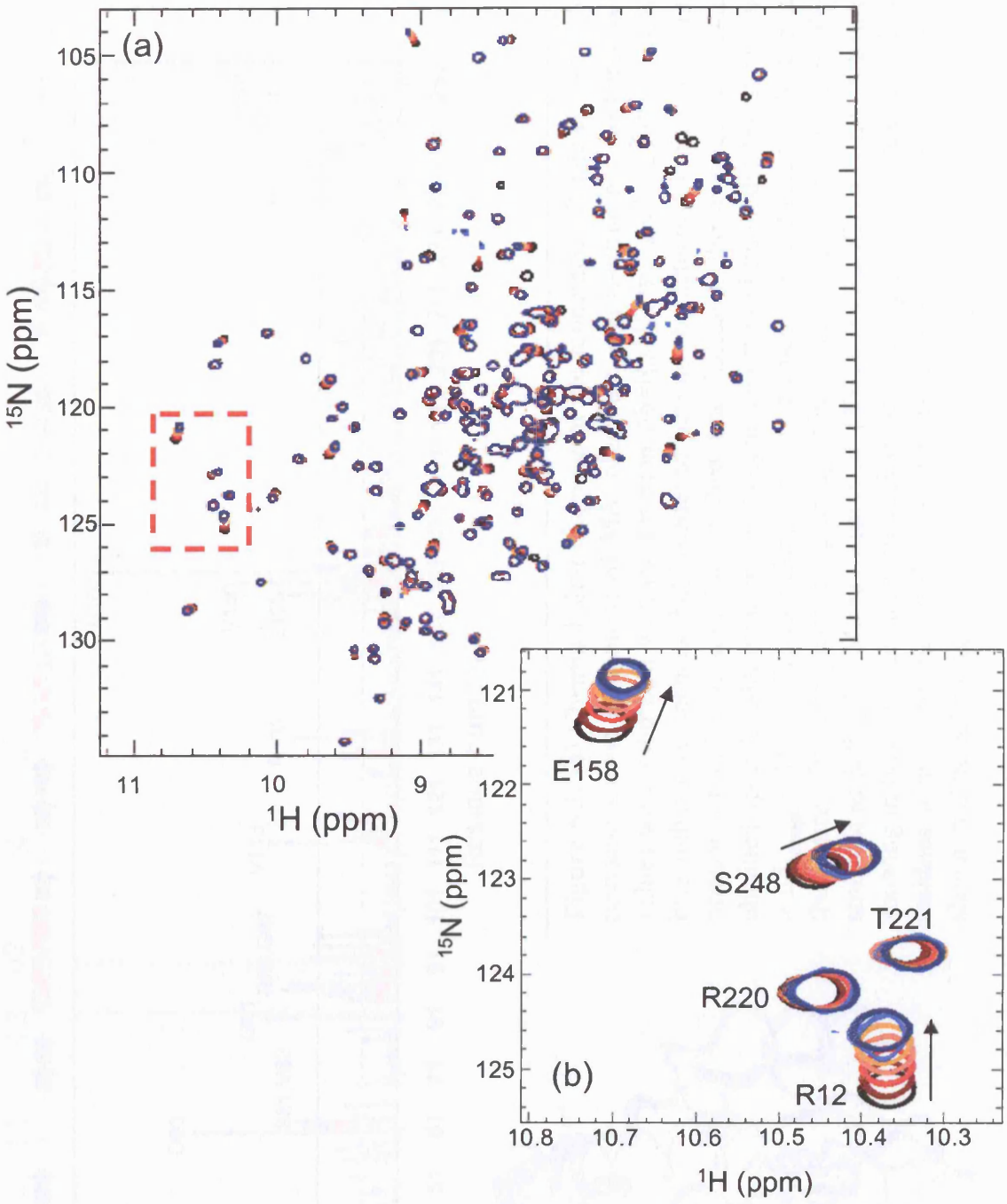

Figure 6.2. (a) Superimposed [^{1}H,^{15}N]-HSQC spectra of the monomeric *Pa*DDAH acquired after the incremental addition of L-citrulline. The black spectrum corresponds to the free protein and the dark blue spectra corresponds to the last titration point upon saturation of the enzyme. (b) Inset of (a) as delimited by the red dotted box. The arrows indicate spectra acquired at increasing L-citrulline concentrations. The spectra were acquired at ^{1}H proton frequency of 500 MHz at 25°C (see Chapter 2, Section 2.5 for details).

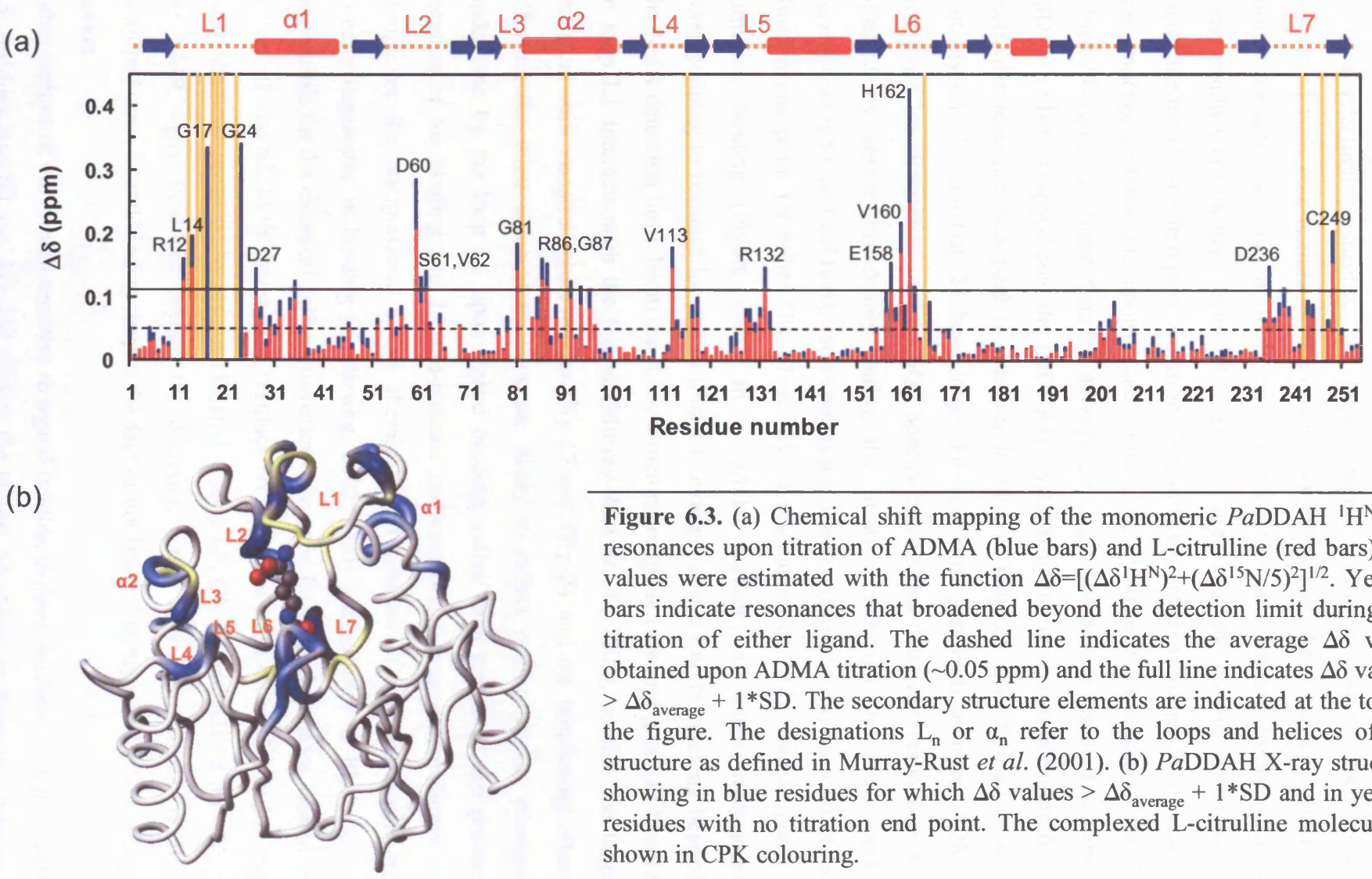

Figure 6.3. (a) Chemical shift mapping of the monomeric *Pa*DDAH $^{1}H^{N}$-^{15}N resonances upon titration of ADMA (blue bars) and L-citrulline (red bars). Δδ values were estimated with the function $\Delta\delta=[(\Delta\delta^{1}H^{N})^{2}+(\Delta\delta^{15}N/5)^{2}]^{1/2}$. Yellow bars indicate resonances that broadened beyond the detection limit during the titration of either ligand. The dashed line indicates the average Δδ value obtained upon ADMA titration (~0.05 ppm) and the full line indicates Δδ values > $\Delta\delta_{average}$ + 1*SD. The secondary structure elements are indicated at the top of the figure. The designations L_n or α_n refer to the loops and helices of the structure as defined in Murray-Rust *et al.* (2001). (b) *Pa*DDAH X-ray structure showing in blue residues for which Δδ values > $\Delta\delta_{average}$ + 1*SD and in yellow residues with no titration end point. The complexed L-citrulline molecule is shown in CPK colouring.

confidently determined due to resonance overlap; these residues included Glu 33, Val 64, Val 68, Glu 83, Glu 171, Val 182, Asn 209 and Glu 210.

From a structural perspective, the largest chemical shift deviations observed are associated with nine regions of the protein. As might have been expected, residues around the active site region are the most sensitive to ligand binding. In broad terms, these residues are located within all seven loop regions of the protein, six of which constitute the active site region. Additionally residues situated in two helices α1 and α2 also undergo perturbations upon ligand binding. To illustrate this result, residues for which $\Delta\delta$ values are larger than the average $\Delta\delta_{average}$ value by one standard deviation (SD) have been mapped onto the *Pa*DDAH crystal structure as shown in Figure 6.3b. Particular attention was paid to residues located in loop L1 over the active site. Notably Gly 17 and Gly 24 have large chemical shift perturbations upon ADMA titration. Close inspection of the HSQC spectra reveals that these two resonances are affected by severe broadening during the titration, although the cross peaks corresponding to the final two titration points are detected (Figure 6.4a). In contrast, no titration end point for either Gly 17 or Gly 24 resonances could be found upon L-citrulline binding (Figure 6.4b). In a similar fashion, all other resonances corresponding to residues located in loop L1 undergo severe line broadening beyond the NMR detection limit. From the X-ray structure, only the carbonyl group of Leu 18 in loop L1 interacts with the ligand (Murray-Rust *et al.*, 2001), so that the large chemical shift variation observed for Gly 17 and Gly 24 and the broadening effect affecting the other adjacent residues are likely to reflect conformational changes undergone by the loop L1 upon ligand binding, rather than report on the precise location of the binding site. Ligand-induced conformational changes are known to complicate the interpretation of the chemical shift changes of residues which are directly implicated in binding (Zuiderweg, 2002). This phenomenon is likely to be responsible for the chemical shift perturbations observed for several residues located in helices α1 and α2. In the latter case, residues 86 and 87 of helix α2 are located in the tilt of the helix preceded by residues located in loop L3, Gly 81 and Ala 82 for which the NMR signal broadens beyond the detection limit. This entire region is also perturbed upon ligand binding despite the fact of not being directly part of the binding pocket.

Other regions of the protein sensitive to ligand binding include residues in loop L2 and L5; residues Asp 60 and Arg 132 exhibit the largest $\Delta\delta$ values in these two regions.

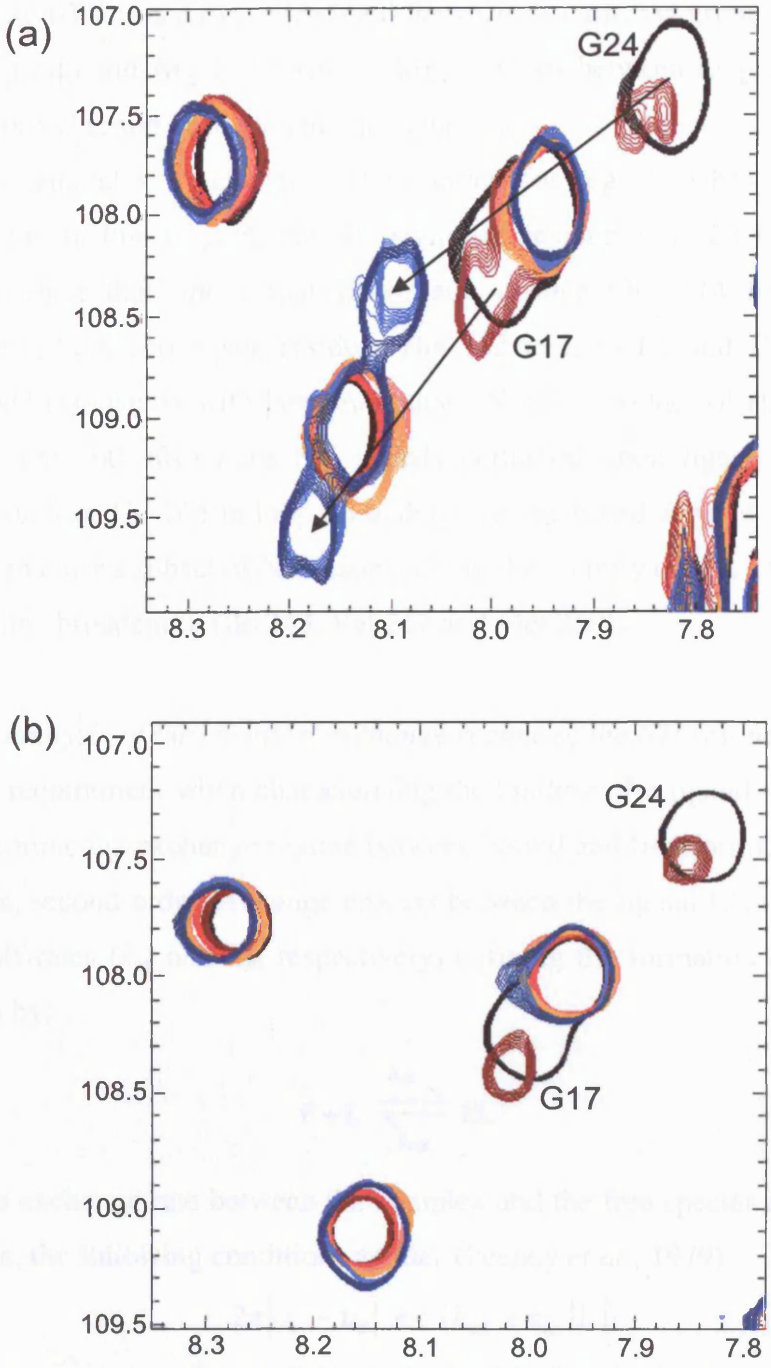

Figure 6.4. Superimposed sections of [^{1}H,^{15}N]-HSQC spectra of monomeric [^{1}H,^{15}N]-labelled *Pa*DDAH acquired upon addition of (a) ADMA and (b) L-citrulline. The black spectrum corresponds to the resonances of the free protein and the dark blue spectrum corresponds to the last titration point upon saturation of the enzyme. The spectra were acquired at ^{1}H proton frequency of 500 MHz at 25°C (see Chapter 2, Section 2.5 for details).

From crystallographic evidence both these residues interact with the bound citrulline ligand. Asp 60 forms a hydrogen bond between its carboxyl group and the L-citrulline α-amide group and Arg 132 forms hydrogen bonds between its guanidino moiety and the α-carboxyl group of the ligand molecule.

Finally, chemical shift changes are observed for regions where the catalytic triad residues lie. In this respect, the NH signal of residue Val 113 is more sensitive to ligand binding than the actual active site residue Glu 114 located in loop L4. Conversely, both active site residues His 162 in loop L6 and Cys 249 in loop L7 exhibit NH resonances with large $\Delta\delta$ values. Nearby residues of His 162, such as Glu 158 and Val 160, also have NH signals perturbed upon ligand binding and other residues such as Thr 165 in loop L6 undergo severe broadening of their line shapes. In a similar manner a subset of NH resonances in the vicinity of Cys 249 (in loop L7) also undergo line broadening (Ile 243, Val 247 and Met 250).

Detailed analysis of the chemical exchange regime of the NH resonances

The first requirement when characterising the binding of a ligand to a macromolecule is to determine the exchange regime between bound and free forms. When considering a two-site, second-order exchange process between the ligand L and the protein P, the on and off-rates (k_{on} and k_{off} respectively) defining the formation of the complex PL are given by:

$$P + L \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} PL \qquad \text{Equation 6.1}$$

When the exchange rate between the complex and the free species is fast on the NMR time scale, the following conditions applies (Feeney *et al.*, 1979):

$$2\pi|\upsilon_F - \upsilon_B| << (k_{off} + k_{on}[L]) \qquad \text{Equation 6.2}$$

where υ_F and υ_B are the resonant frequencies (Hz) of the nucleus in the free and the bound states, [L] is the ligand concentration. A single resonance is observed in this case at the frequency defined as $\upsilon = p_F\upsilon_F - p_B\upsilon_B$, where p_F and p_B are the fractional populations of the free and bound states. Under these conditions, the dissociation binding constant for this reaction is then defined as:

$$K_d = \frac{[P][L]}{[PL]} = \frac{k_{off}}{k_{on}} \qquad \text{Equation 6.3}$$

Hence, a distinct advantage of screening compounds by NMR is the possibility of determining the dissociation constants for the complex. At equilibrium, [PL] = pP_0, [P] = (1-p)[P_0], and [L] = [L_0] –p[P_0], where [L_0] and [P_0] are the total concentrations of ligand and protein respectively. These equations can be substituted in Equation 6.3 to yield:

$$p = \frac{([P_0]+[L_0]+[K_d]-\sqrt{([P_0]+[L_0]+[K_d])^2-4[P_0][L_0]}}{2[P_0]} \qquad \text{Equation 6.4}$$

Where the parameter p represents the fractional populations of bound and free states at equilibrium (Pellecchia, 2005): $p = (\delta_{obs} - \delta_{free})/(\delta_{sat} - \delta_{free})$, δ_{obs} is the observed chemical shift measured during the titration, and δ_{free} and δ_{sat} are the chemical shifts of the free state and the fully bound state of the protein respectively. Hence, the chemical shift mapping by stepwise titration of the ligand can provide an estimate of K_d through a nonlinear least squares fit of p as a function of total ligand concentration [L_0] or, in some instances, of the molar ratio between ligand and protein concentrations (see Chapter 2, Section 2.5 for more details).

Alternatively, in the slow exchange regime, when the life times of the bound and free forms of the protein are long relative to the chemical shift difference time scale (k_{off} + k_{on}[L]) << $2\pi\Delta\upsilon$, two resonant frequencies are ultimately observed. The binding constant can still be determined by measuring the intensities of the resonances from the disappearing peak and/or the appearing peaks as a function of the concentration of added ligand.

However, in the intermediate exchange regime, when (k_{off} + k_{on}[L]) ≈ $2\pi\Delta\upsilon$, differential line broadening precludes the estimation of a dissociation constant as the relevant resonances often disappear from the 2D NMR spectra.

In the case of *Pa*DDAH, the titration of either ADMA or L-citrulline did not reveal any resonance with behaviour consistent with the slow exchange regime. However, several NH resonances either shift progressively without change in intensity, which is characteristic of the fast exchange regime, or experience just slight line broadening at intermediate ligand concentrations. In the latter case, the resonances were considered as being in a "moderately fast" exchange rate regime as defined by Lian and Roberts,

(1993). In practice, in cases where only slight exchange contributions to the line widths are observed for a particular resonance, the fast exchange regime is still assumed to facilitate the analysis. In this vein, the amide groups of Cys 249 and Gly 87 were assumed to be under time fast exchange kinetics in order to estimate qualitatively a dissociation constant for the complexes (Figure 6.5). The normalised chemical shift changes of these cross peaks in either ^{1}H and ^{15}N dimensions were plotted as a function of the molar ratio of ligand to protein. The molar ratio was chosen instead of the actual ligand concentration in Equation 6.4 to account for protein dilution as the titration proceeded (see Chapter 2, Section 2.5). The nonlinear least-squares fit to the data using Equation 6.4 yielded the binding curves for ADMA and L-citrulline shown in Figure 6.6. From the chemical shift changes obtained from Gly 87 and Cys 249, ADMA has a mean estimated dissociation constant of 413 ± 63 μM (n = 2). The K_d value obtained for ADMA is consistent with the K_M value of this enzyme for the substrate which is around 460 μM. In contrast, L-citrulline had a mean estimated 8-fold lower affinity for *Pa*DDAH as the K_d obtained is 3.35 ± 0.57 mM (n = 2). This result is also consistent with findings that indicate L-citrulline does bind to the active site at very high concentrations (Murray-Rust *et al.*, 2001).

However, several amide resonances broaden markedly as ligand concentration increases becoming undetectable. The signal reappears at the highest ligand concentrations and can still shift as the concentration is further increased. This behaviour is characteristic of the intermediate exchange regime (Lian and Roberts, 1993). In particular, resonances that display this behaviour correspond to residues involved in direct interaction with the ligand, as exemplified by the effect of ADMA titration on Asp 60 and His 162 shifting resonances (Figure 6.7). For resonances in the intermediate exchange regime, the approximation $(k_{off} + k_{on}[L]) \approx 2\pi\Delta\upsilon$ can relate the broadening of the cross peaks of the free and bound forms of the protein to the rate constant k_{off} according to $k_{off} \approx 2\pi\Delta\upsilon$, where $k_{on}[L]$ is assumed to be small relative to k_{off} (Katoh *et al.*, 2003).

Under this assumption, a qualitative estimate of the k_{off} rates from the chemical shift changes of the NH resonances of Asp 60 and His 162 was obtained. On average the amide proton of these residues shifts 7-44 Hz in the $^{1}H^{N}$ dimension which corresponds to a dissociation rate k_{off} of 44 to 276 s^{-1}. Similarly, several amide cross peaks corresponding to residues in loop L1 also experience severe line broadening midway

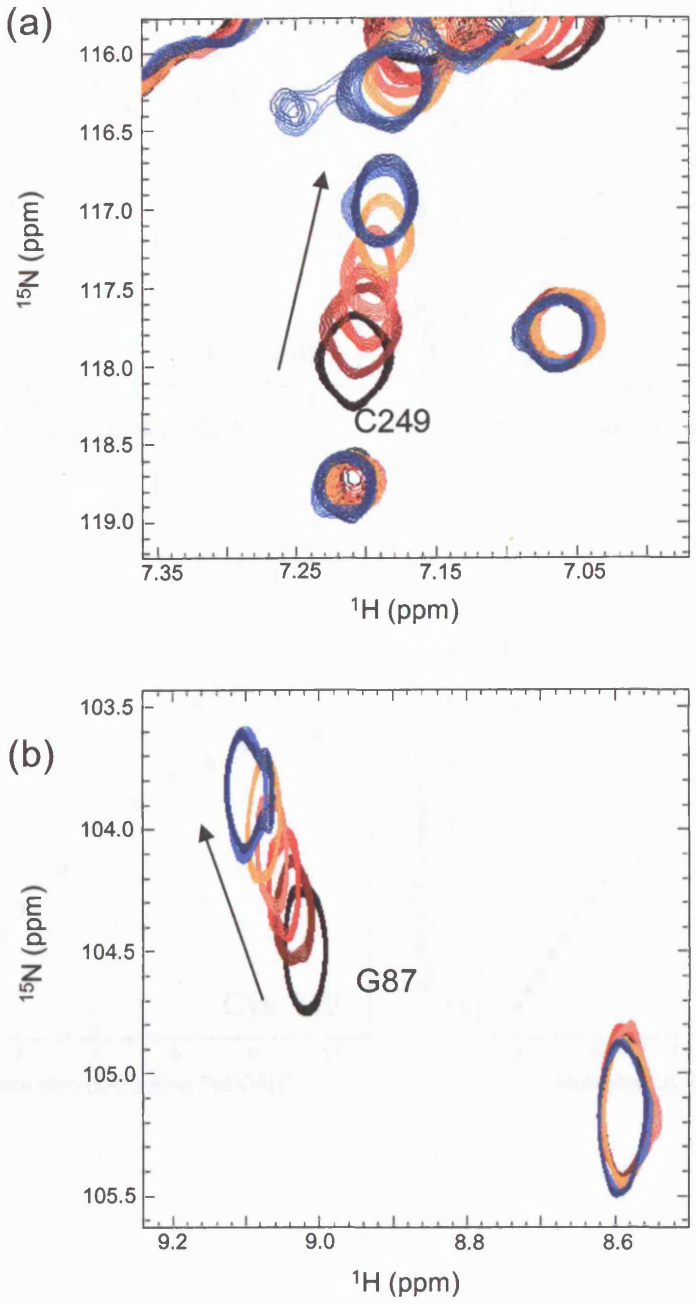

Figure 6.5. Examples of HN resonances under the time fast to "moderately fast" exchange regimes. Superimposed sections of [^{1}H,^{15}N]-HSQC spectra of monomeric [^{1}H,^{15}N]-labelled *Pa*DDAH showing the resonances corresponding to residues (a) Cys 249 and (b) Gly 87 upon the incremental addition of ADMA. The black spectrum corresponds to the free protein resonances and the dark blue spectrum corresponds to the last titration point. The spectra were acquired at ^{1}H proton frequency of 500 MHz at 25°C (see Chapter 2, Section 2.5 for details).

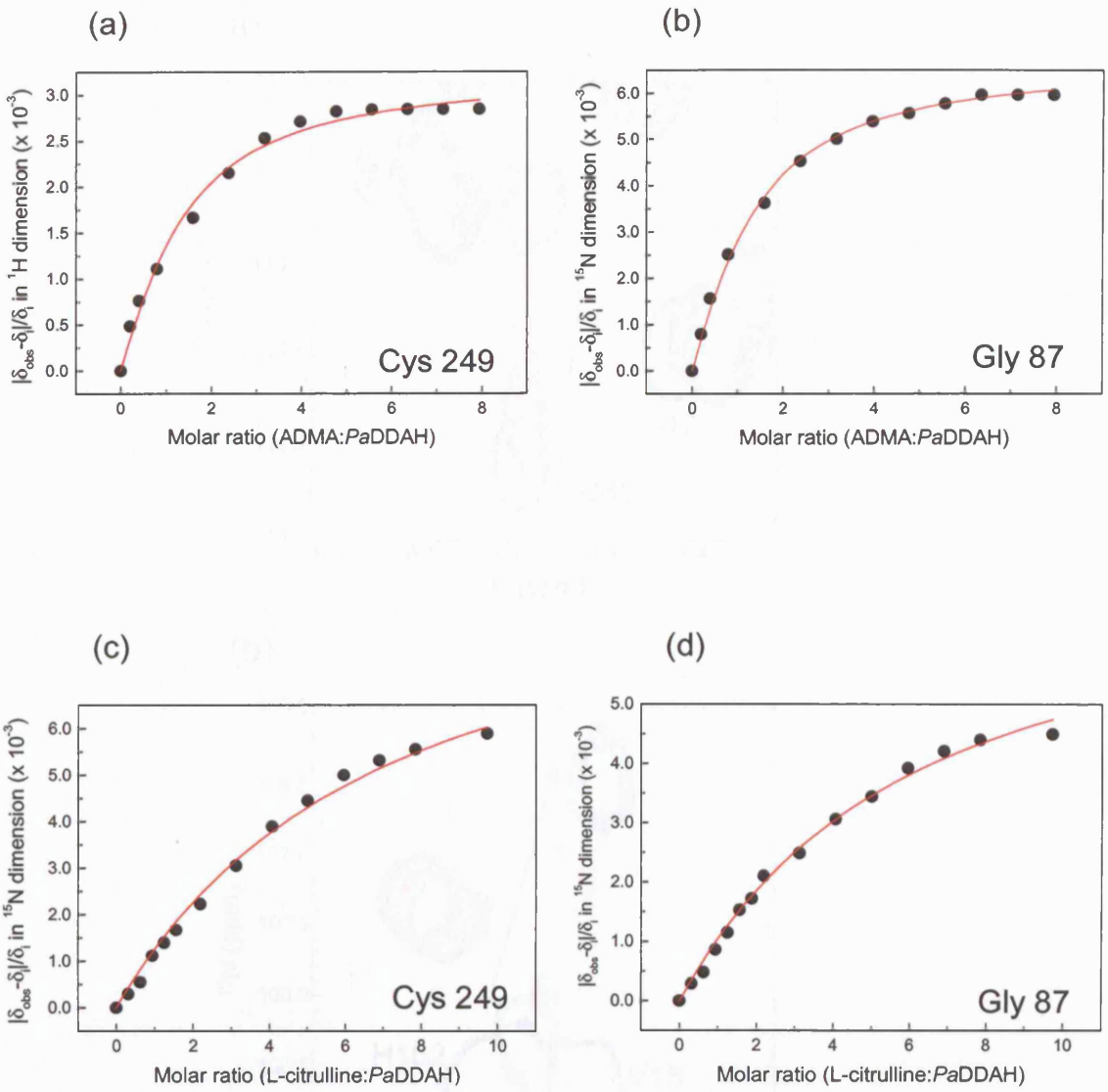

Figure 6.6. Examples of binding-curves obtained for residues Gly 87 and Cys 249 following the titration of ADMA (a,b) and L-citrulline (c,d) (see Chapter 2, Section 2.5 for details). The curves were obtained from a non-least squared fitting of the data to equation 6.4.

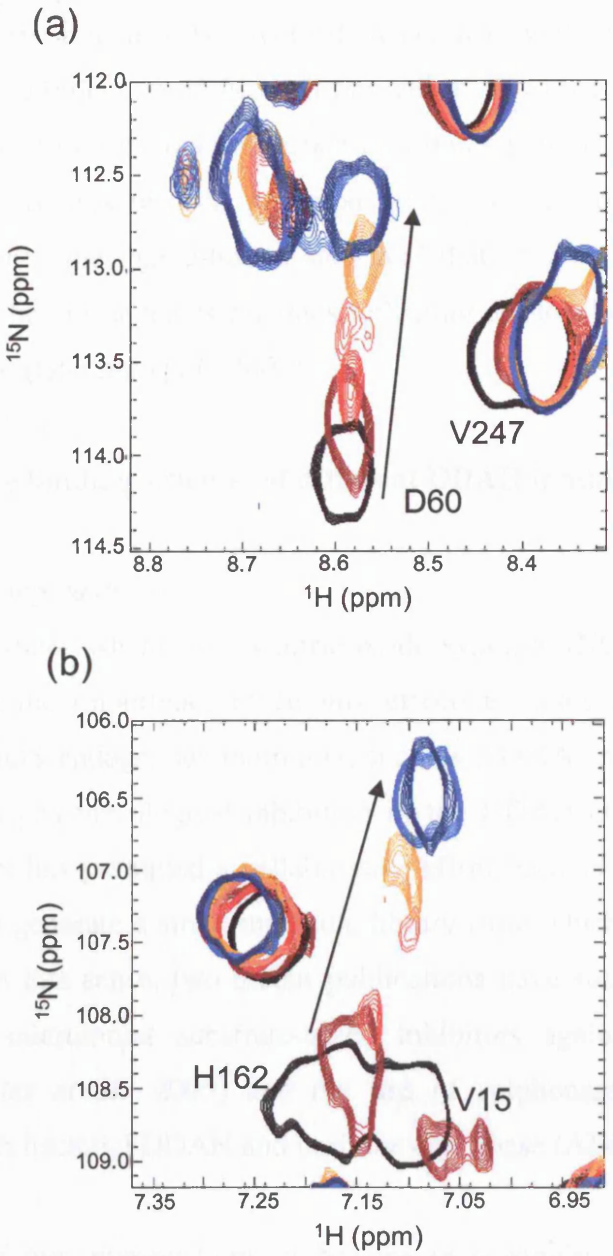

Figure 6.7. Examples of HN resonances under the time intermediate exchange regime. Superimposed sections of [^{1}H,^{15}N]-HSQC spectra of monomeric [^{1}H,^{15}N]-labelled *Pa*DDAH showing the resonances corresponding to residues (a) Asp 60 and (b) His 162 upon the incremental addition of ADMA. The black spectrum corresponds to the free protein resonances and the dark blue spectrum corresponds to the last titration point. The spectra were acquired at ^{1}H proton frequency of 500 MHz at 25°C (see Chapter 2, Section 2.5 for details).

through the titration of ADMA (Figure 6.4a). Higher k_{off} values around 600 and 860 s^{-1} were estimated for residues Gly 17 and Gly 24, respectively (Figure 6.4a). Since Gly 17 and Gly 24 are not directly involved in contacts with the ligand, the k_{off} rates estimated are not a reflection of the local on- and off-rates between ligand and protein, but rather they report on the overall response of this region of the protein to the ligand binding process. In this respect, the apparent k_{off} values might reflect the rate of conversion between the ligand-bound state and that of apo-form of the protein and thus, they could be envisaged as off-rates reflecting the local dynamics induced upon ligand interaction (Günther *et al.*, 2002).

6.3 Assessing the binding schemes of different DDAH inhibitors

The classes of compounds

Several disease states where excess nitric oxide synthase (NOS) activity is involved have underlined the importance of finding effective routes to increase the *in vivo* availability of NOS endogenous inhibitors, such as ADMA and L-NMMA. One such route entails the pharmacological inhibition of the DDAH enzymes. The search for DDAH inhibitors has prompted a collaborative effort between several departments at UCL in order to generate a small molecule library from which these inhibitors might be developed. In this sense, two recent publications have successfully described the design of low micromolar substrate-based inhibitors against mammalian DDAH enzymes (Rossiter *et al.*, 2005) and the use of sulphonamide derivatives in the inhibition of both bacterial DDAH and arginine deiminase (ADI) enzymes (Vallance *et al.*, 2005).

In the course of our investigations, it became increasingly clear that bacterial and mammalian DDAH enzymes have different inhibitor profiles. The screening of substrate-based inhibitors described as competitive inhibitors of mammalian DDAH enzymes failed to produce any real hit against *Pa*DDAH. One molecule however, SR 140 (Figure 6.8A) did show low micromolar inhibitory concentration IC_{50} ~ 30 μM against *Pa*DDAH (Dr R. Hurtado, *personal communication*) and was consequently chosen as a candidate for NMR titration experiments. Additionally, SR 227 and SR 257, described by Rossiter *et al.* (2005) as being good mammalian DDAH I inhibitors, were further scrutinised by our NMR experiments despite their rather high IC_{50} values measured against *Pa*DDAH (>1 mM; Dr R. Hurtado, *personal communication*; Figure

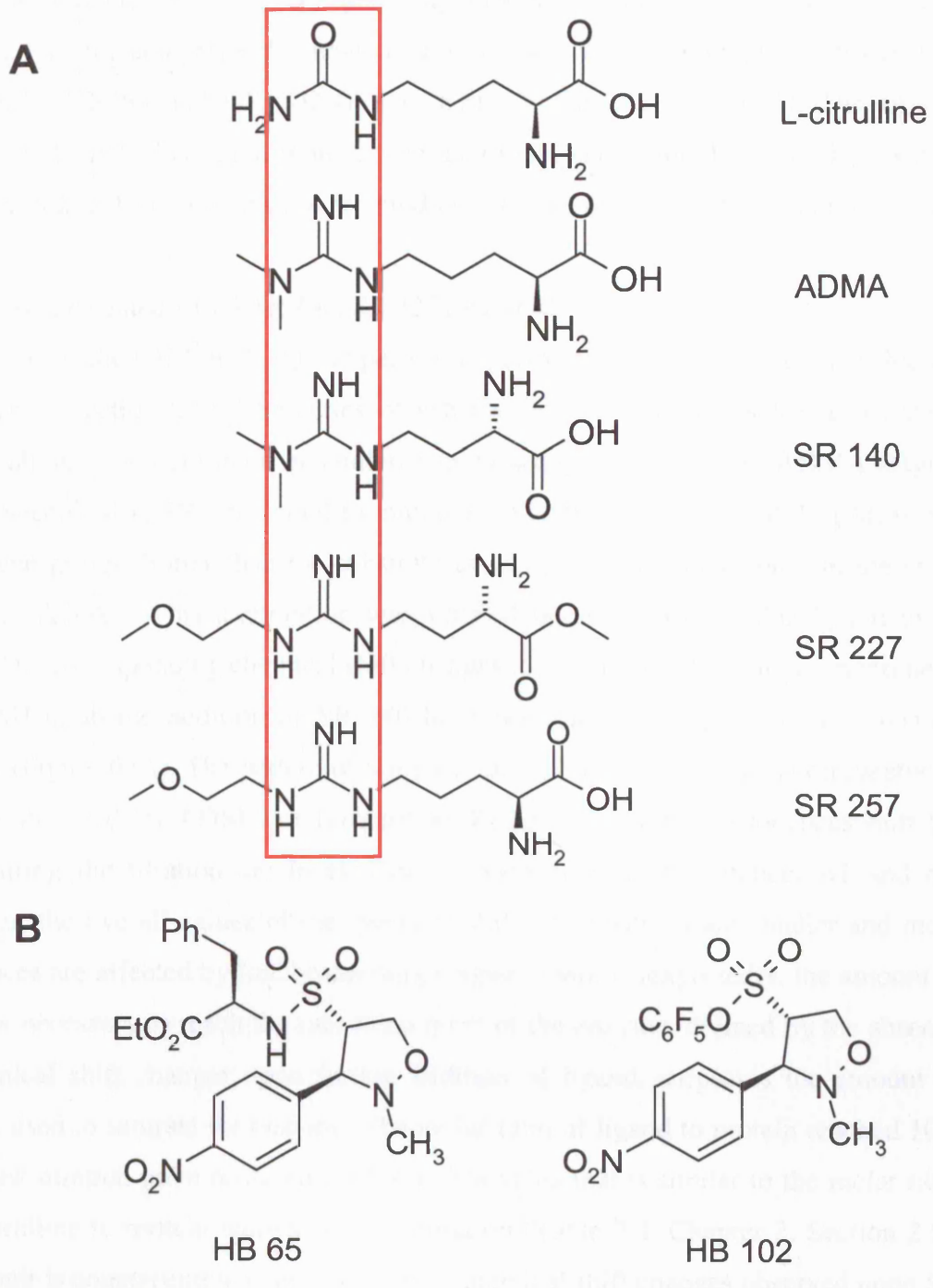

Figure 6.8. Molecular structures of the different compounds used in the NMR titration experiments. (A) The red box highlights the conservation of the guanidino moiety in substrate-based compounds. (B) HB 65 and HB 102 are sulphonamide and pentafluorophenyl sulfonate compounds respectively.

6.8A). Moreover, the discovery that sulphonamides could inhibit *Pa*DDAH also prompted us to characterise this new class of molecules with two representatives from the family, HB 65 and HB 102 (Figure 6.8B; Vallance *et al.*, 2005). The authors suggest that *Pa*DDAH could represent an additional antibacterial target which would further validate the comparison of the binding schemes of these different compounds.

Titrations experiments with SR 140, SR 227 and SR 257

The details of the [^{1}H,^{15}N]-HSQC experiments performed are summarised in Table 2.4 (Chapter 2, Section 2.5). The series of substrate-based inhibitors selected for these studies all share the common guanidino moiety seen in the substrate ADMA (Figure 6.8A). Specifically, SR 140 found to inhibit *Pa*DDAH with an IC_{50} of 30 μM, is one methylene group shorter than the substrate and in principle should bind in the same cavity as ADMA. This assumption was verified upon titration of this ligand using NMR. The corresponding chemical shift changes in the HSQC spectrum of monomeric *Pa*DDAH upon the addition of SR 140 have been plotted as a function of residue number (Figure 6.9a). The pattern of chemical shift changes obtained is comparable to the one observed for ADMA or L-citrulline. Residues for which resonances shift the most during the titration are located in the loop areas and in helices α1 and α2. However, the overall values of the chemical shift perturbations are smaller and more resonances are affected by line broadening (Figure 6.9a). Unexpectedly, the amount of inhibitor necessary to reach the saturation point of the enzyme, defined by the absence of chemical shift changes upon further addition of ligand, surpasses the amount of ADMA used to saturate the enzyme. The molar ratio of ligand to protein reached 10:1 at the last titration point acquired for SR 140, a value that is similar to the molar ratio of L-citrulline to protein reached upon saturation (Table 2.4, Chapter 2, Section 2.5). This result is counterintuitive since the small chemical shift changes observed upon SR 140 titration and the increased degree of line broadening observed in comparison with the results obtained with ADMA suggest that SR 140 is even a weaker binder than ADMA. At this stage, it is unclear whether the derived IC_{50} value by enzyme essay is a valid indication of the overall binding affinity towards the enzyme. A K_d value could not be derived from the NMR titration as the chemical shift changes of residue Cys 249 or Gly 87 could not be fit satisfactorily to Equation 6.4.

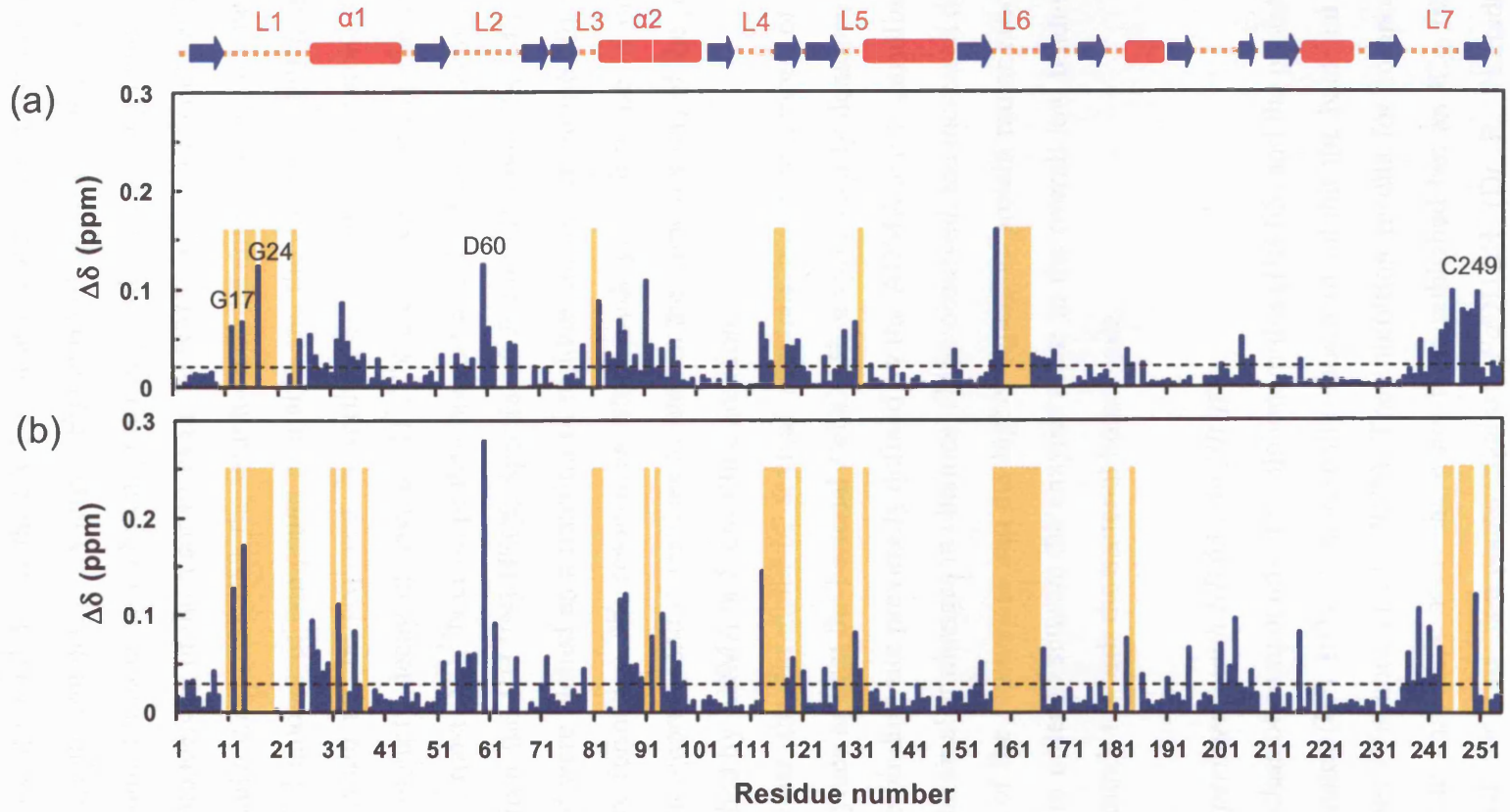

Figure 6.9. (a) Chemical shift mapping of the monomeric *Pa*DDAH $^{1}H^{N}$-^{15}N resonances upon titration of SR 140 and (b) SR 227. $\Delta\delta$ values were estimated with the function $\Delta\delta=[(\Delta\delta^{1}H^{N})^{2}+(\Delta\delta^{15}N/5)^{2}]^{1/2}$. Yellow bars indicate resonances that broadened beyond the detection limit during the titration of either ligand (the scaling along the y-axis is arbitrary). The dashed line indicates the average $\Delta\delta$ value obtained upon titration. The secondary structure elements are indicated at the top of the figure. The designations L_n or α_n refer to the loops and helices of the structure as defined in Murray-Rust *et al.* (2001).

With respect to the titration of the ether-based compound SR 227, the number of *Pa*DDAH resonances affected by exchange broadening again increased in comparison to the number observed during the SR 140 titration, although the observed changes in chemical shift are comparatively larger (Figure 6.9b). However SR 227 was not a good *Pa*DDAH inhibitor as assessed by enzyme assay and it is tempting to speculate that the additional broadening of the resonances in the NMR experiment could reflect this fact. In the particular case of SR 257, the saturation of the enzyme could not be reached due to the limited amount of compound available and the chemical shift mapping of this compound could not be achieved. Nevertheless, in order to illustrate the recurring pattern of intensity decay of several $^{1}H^{N}$,^{15}N cross peaks during the titrations, the intensities of most $^{1}H^{N}$,^{15}N cross peaks obtained at a 1:1 ligand-protein complex were extracted from the [^{1}H,^{15}N]-HSQC spectra of *Pa*DDAH upon SR 140 and SR 257 addition and were plotted as a function of residue number (Figure 6.10). Clearly, at a 1:1 complex formation, the resonances which undergo a decrease in intensity upon ligand titration correspond to residues located in the same regions of the protein which were perturbed by ADMA or L-citrulline addition.

Taken together the titrations of several substrate-based inhibitors of mammalian DDAH enzymes against the bacterial *Pa*DDAH enzyme clearly demonstrate that they can map the binding site previously defined by the ADMA and L-citrulline molecules. However, the smaller changes in chemical shifts observed, the increased degree of line broadening of the resonances and the higher ligand-to-protein molar ratio that had to be reached in order to saturate the enzyme attest to the overall low binding affinity of these compounds towards the bacterial homologue.

Titration experiments with HB 65 and HB 102

A different class of compounds, the sulphonamides (HB 65) and the pentafluorophenyl (PFP) sulfonates (HB 102), were recently shown to inhibit the bacterial DDAH and ADI enzymes (Vallance *et al.*, 2005). The inhibition results for compound HB 65, based upon an enzymatic assay, have not been published but an IC_{50} of 20 μM has been derived for this compound against *Pa*DDAH (Dr R. Hurtado, *personal communication*). On the other hand, HB 102 refers to compound 3 in Vallance *et al.* (2005) and has an IC_{50} of 21 μM against *Pa*DDAH. Due to their low solubility in aqueous buffers, both compounds were dissolved in dimethyl sulfoxide (DMSO) and could only be titrated into *Pa*DDAH up to a final molar ratio of 3:1 ligand to protein.

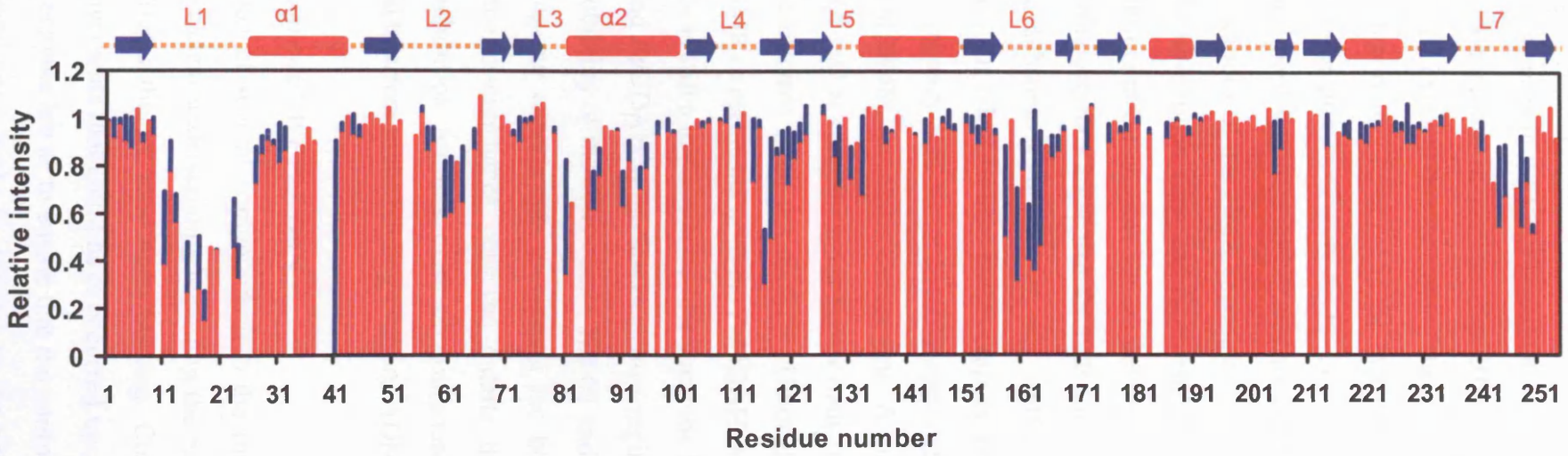

Figure 6.10. Normalised intensities of the $^{1}H^{N}$,^{15}N correlations extracted from the 2D-HSQC spectra of the 1:1 complex of SR 140 with *Pa*DDAH in red and SR 257 with *Pa*DDAH in blue. The secondary structure elements are indicated at the top of the figure. The designations L_n or α_n refer to the loops and helices of the structure as defined in Murray-Rust *et al.* (2001).

Additionally, a control experiment was performed in order to determine whether DMSO would bind to the protein. In fact, DMSO alone does alter the chemical shift pattern of two resonances in the [^{1}H,^{15}N]-HSQC spectra of *Pa*DDAH corresponding to the $^{1}H^{N\varepsilon}$ side chain proton of residue Trp 207 and the amide group of Gly 229. Presumably a pocket exists at this location of the enzyme to accommodate such a small molecule. More importantly however is the fact that neither compound perturbed the chemical shifts of the protein (Figure 6.11). In contrast to the titrations of substrate-based compounds, these sulfonate derivatives are not targeted to the active site, at least not under the conditions of our NMR experiments. The authors describe HB 102 as a potent *Pa*DDAH inhibitor *in vitro* and have established that the PFP derivatives are competitive inhibitors of the enzyme but are not clear as to their precise mechanism of inhibition. Nevertheless, they argue that the PFP-sulfonates must bind the protein more tightly than the substrate (Vallance *et al.*, 2005). This suggestion is clearly in disagreement with our NMR results. In this respect, possible considerations aiming to explain such discrepancy will be necessarily speculative. One concern lies with the potential ability of DMSO to inhibit the enzyme. A qualitative assay did confirm that the presence of DMSO at concentrations higher than 10% (v/v) in the reaction mixture could inhibit the enzyme (data not shown). Although this concentration was never reached in our NMR experiments or in the *Pa*DDAH inhibition assays presented by the authors, this issue is still a matter to consider since the NMR titrations have shown that DMSO could bind *Pa*DDAH in the fast exchange regime (Figure 6.11). Alternatively, due to the low solubility of the compounds, HB 65 and HB 102 could have aggregated in solution during our experiments precluding the binding to the enzyme. Another possibility for the discrepancies observed entails the existence of an alternative mechanism of inhibition for this class of compounds, in such a way that these compounds could interfere in the binding event of ADMA.

Control experiments with WT PaDDAH

The NMR titration of several compounds into the monomeric *Pa*DDAH has so far failed to provide a clear understanding regarding the exact basis of the conformational exchange that affects the L1 loop upon binding. Concerns over the impact of the structural rearrangements that could have occurred upon the introduction of two point mutations in the enzyme led us to ensure that the conformational exchange broadening observed for several cross peaks is intrinsic to the bound state of the enzyme rather

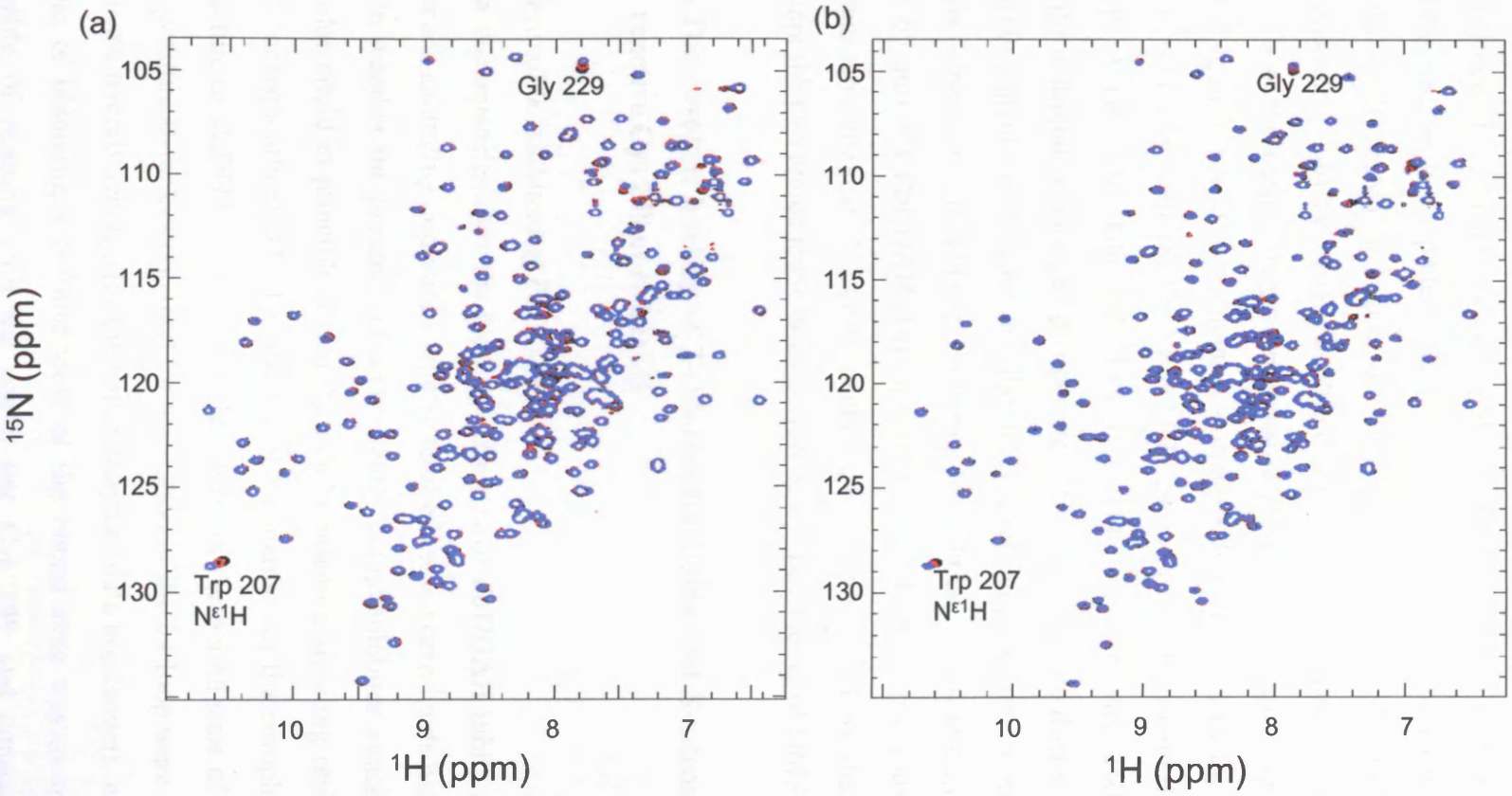

Figure 6.11. (a) Superimposed [^{1}H,^{15}N]-HSQC spectra of the monomeric [^{1}H,^{15}N]-labelled *Pa*DDAH acquired after the incremental addition of HB 65 (a) and HB 102 (b). The black spectra correspond to the free protein, the red spectra to that of a 1:1 ligand-protein complex and the light blue spectra correspond to a 3:1 ligand-protein complex. The spectra were acquired at ^{1}H proton frequency of 500 MHz at 25°C (see Chapter 2, Section 2.5 for details).

than an artefact of the monomerisation strategy. In this vein, SR 140 and HB 65 were titrated into the ^{15}N-labelled WT *Pa*DDAH protein. The results obtained were analysed by inspecting changes in chemical shift or cross peak intensity in the [^{1}H,^{15}N]-TROSY spectra. By reference to the titration studies of the monomeric enzyme, the pattern of broadened $^{1}H^{N}$-^{15}N resonances of WT *Pa*DDAH observed upon titration of SR 140 agrees well with the pattern of cross peaks affected in the monomeric enzyme. This result is illustrated in Figure 6.12 by several [^{1}H,^{15}N]-TROSY sections in which resonances could be easily identified through the comparison with [^{1}H,^{15}N]-HSQC spectra of the double mutant monomeric protein at a ligand-to-protein molar ratio of 2:1 (Figure 6.12). The resonances corresponding to residues located in the L1 loop, such as Gly 14, Leu 18, Thr 19 and Gly 24 and residues around the active site regions such as Gly 116, Glu 158, Lys 164, Thr 165 and Met 250 are all subject to conformational exchange broadening. This finding confirms that the structural rearrangements undergone by the loop upon binding are reproduced in both the homodimeric and monomeric proteins. In addition, the titration of the sulphonamide HB 65 into WT *Pa*DDAH also supports the NMR results previously obtained with the monomeric enzyme. No cross peaks have been affected by the addition of HB 65 except the two amide resonances perturbed by the binding of DMSO (data not shown).

6.4 The covalent binding of 2-chloroacetamidine and S-nitroso-L-cysteine to the reactive Cys 249 of *Pa*DDAH

Irreversible inhibitors of PaDDAH

The discrepancies over the binding affinities of *Pa*DDAH inhibitors and the inability to reach saturating conditions with several of these compounds has raised the question as to whether the presence of a covalently bound inhibitor attached to an active site residue could in principle lessen the line broadening affecting residues in the L1 loop and consequently enable the study of the dynamics of the complex. Since the crystal structure of *Pa*DDAH provided a clear view of the bound state of the enzyme and the temperature B-factors obtained for the region of the L1 loop were not evidently higher than the overall average for the entire enzyme (data not shown), we surmised that one route of obtaining a definite view of the bound state was to search for molecules capable of reacting with the active site Cys 249 and consequently locking the conformation of the enzyme into the bound form.

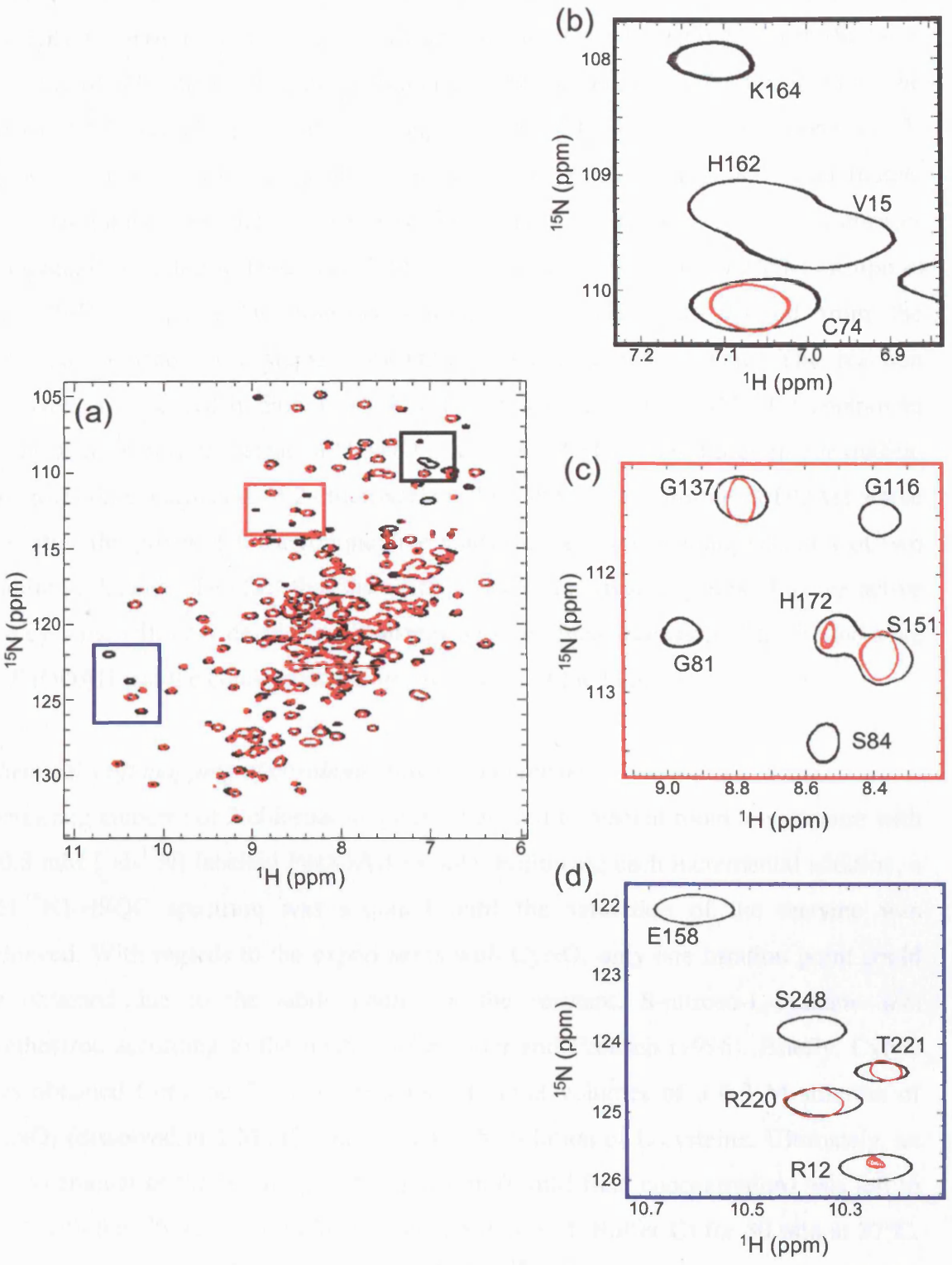

Figure 6.12. (a) Superimposed [^{1}H,^{15}N]-TROSY spectra acquired of the [^{1}H,^{15}N]-labelled WT *Pa*DDAH in the absence (black spectra) and presence of the ligand SR 140 (red spectra) at a molar ratio of 2:1 ligand-to-protein. (b-d) Insets of (a) showing sections of the spectra were resonances undergo line broadening. The spectra were acquired at ^{1}H proton frequency of 500 MHz at 25°C (see Chapter 2, Section 2.5 for details).

Recently, the small molecule 2-chloroacetamidine was shown to irreversibly inhibit not only *Pa*DDAH but also the peptidylarginine deiminase (PAD) enzymes (Stone *et al.*, 2005b). The mechanism of inhibition is shown in Figure 6.13a and involves the nucleophilic attack of the thiol group of Cys 249 on the C-Cl bond of 2-chloroacetamidine. More interestingly, Knipp and co-workers suggested an alternative route to inhibit the DDAH enzymes. They have synthesised reactive S-nitroso-compounds, which covalently attach to the active site cysteine of DDAH I (Knipp *et al.*, 2005). In particular, S-nitroso-L-homocysteine (HcyNO) derived from the precursor molecule L-cysteine irreversibly inactivated the enzyme. The reaction involved is displayed in Figure 6.13b with S-nitroso-L-cysteine (CyNO) compound acting as the substrate instead of HcyNO. Ultimately, CyNO was chosen in our studies. The titration experiments of 2-chloroacetamidine and CyNO into the *Pa*DDAH active site offer the prospect for a comparative study between the binding schemes of two additional ligands. The fact that these compounds are covalently linked to the active site cysteine will provide additional insight into the characterisation of the bound-state of *Pa*DDAH and the conformational heterogeneity of the loop L1.

Chemical shift mapping of covalently linked compounds

Increasing amounts of 2-chloroacetamidine were left to react at room temperature with a 0.8 mM [^{1}H-^{15}N]-labelled *Pa*DDAH sample. Following each incremental addition, a [^{1}H,^{15}N]-HSQC spectrum was acquired until the saturation of the enzyme was achieved. With regards to the experiments with CyNO, only one titration point could be obtained due to the labile nature of the reactant. S-nitroso-L-cysteine was synthesized according to the method of Stamler and Feelisch (1996). Briefly, CyNO was obtained from the 5 minute reaction of equal volumes of a 0.3 M solution of $NaNO_2$ (dissolved in 1 M HCl) and of a 0.3 M solution of L-cysteine. Ultimately, an excess amount of the resulting CyNO solution (6 mM final concentration) was left to react with the ^{15}N-labelled *Pa*DDAH sample (0.3 mM, Buffer C) for 30 min at 37°C. The subsequent acquisition of several [^{1}H,^{15}N]-HSQC spectra confirmed the completion of the reaction as no further changes in the spectra could be observed. Due to the lack of intermediate titration points, several amide resonances could not be confidently re-assigned in the ligand bound state of the protein (in this case, residues 31, 57, 58, 123, 126, 157, 238, 249).

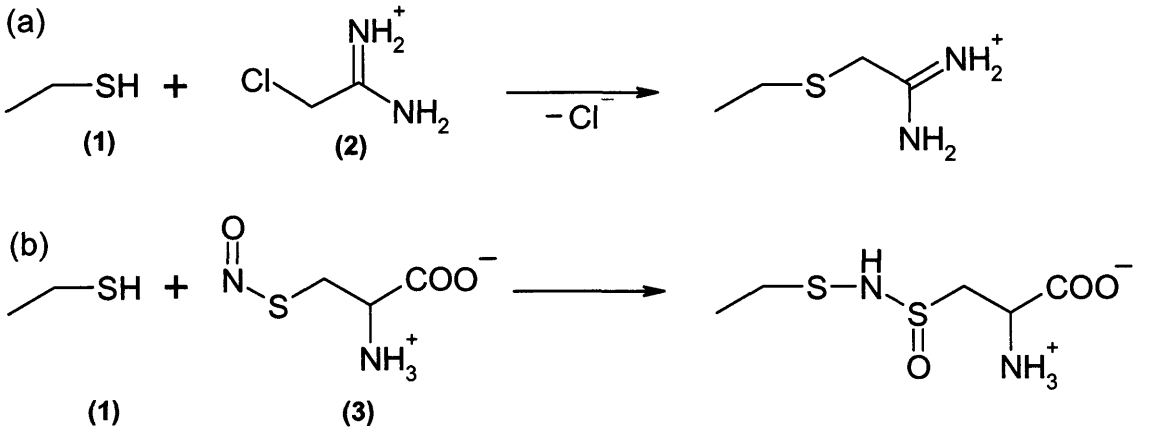

Figure 6.13. Nucleophilic attack of the reactive Cys 249 **(1)** of *Pa*DDAH against (a) 2-chloroacetamidine **(2)** and (b) S-nitroso-L-cysteine **(3)**.

The mapping of the chemical shift changes induced upon the titration of the molecules is plotted as a function of residue number in Figure 6.14. The overall magnitude of the perturbations induced by 2-chloroacetamidine is comparable to the changes induced by L-citrulline (Figure 6.14a). However, contrary to what was observed for the substrate and product molecules, only four main regions of the protein are significantly perturbed by 2-chloroacetamidine. These regions are confined to residues Gly 81 in loop L3 at the entrance of the active site; to residues Ala 115 and Gly 116 in loop L4 where the active site residue Glu 114 lies; to residues 158-160 and Leu 163 located in loop L6 with the additional line broadening of resonances belonging to Leu 161 and His 162; and to residues in loop L7 where the third active site residue Cys 249 is situated. Interestingly, 2-chloroacetamidine binding did not interfere with the chemical exchange kinetics of the amide groups belonging to residues located in loop L1 and NH resonances corresponding to residues could easily be assigned in the bound state. This finding is in contrast with what has been observed upon ADMA or L-citrulline titration which invariably led to the almost complete broadening of the resonances in question.

In comparison, the reaction with CyNO, which is a larger molecule relative to 2-chloroacetamidine, induces the pattern previously observed of $^{1}H^{N}$-^{15}N resonances which "disappear" for residues in the L1 loop region (Figure 6.14b, residues represented by yellow bars). When discussing the titration results of ADMA, it was argued that the broadening of the resonances of the L1 loop reflected some degree of conformational fluctuation of this region induced upon ligand binding. The fact that CyNO is attached to the cysteine residue in the active site disallows the notion that the k_{off} rates of CyNO binding could have been responsible for the line broadening effect. In this respect, the interpretation of these data suggests that the bound state of *Pa*DDAH is not defined by a single conformation and that the loop L1 over the active site is not locked into a single position as suggested by the X-ray structure. The L1 loop might sample multiple conformations in the ligand-bound form of the enzyme with an altered motional regime with respect to the apo-form of *Pa*DDAH. This movement seems independent of the reversibility of ligand binding but is apparently modulated by the size of the molecule occupying the binding pocket.

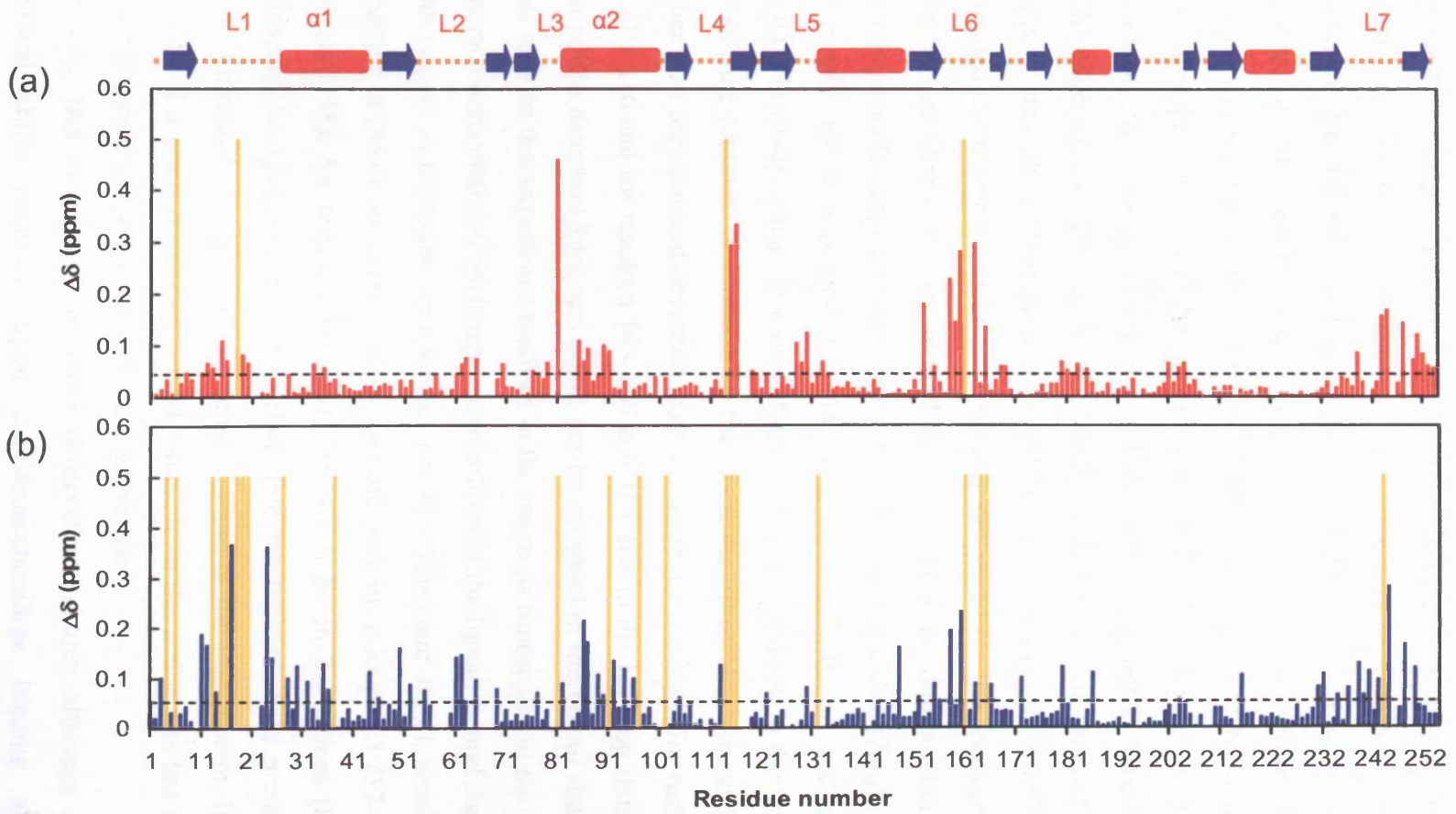

Figure 6.14. Chemical shift mapping of the monomeric *Pa*DDAH $^{1}H^{N}$-^{15}N resonances upon the reaction with (a) 2-chloroacetamidine (red) and (b) S-nitroso-L-cysteine (blue). Δδ values were estimated with the function $\Delta\delta=[(\Delta\delta^{1}H^{N})^{2}+(\Delta\delta^{15}N/5)^{2}]^{1/2}$. Yellow bars indicate resonances that broadened beyond the detection limit (the scaling along the y-axis is arbitrary). The full line indicates the average Δδ value obtained upon titration. The secondary structure elements are indicated at the top of the figure. The designations L_n or α_n refer to the loops and helices of the structure as defined in Murray-Rust *et al.* (2001).

^{15}N relaxation studies of the monomeric PaDDAH bound to 2-chloroacetamidine

The titration of 2-chloroacetamidine into the monomeric *Pa*DDAH active site showed that this small compound did not significantly affect the characteristic ligand-dependent exchange broadening of the amide NH cross peaks assigned to residues in the L1 loop. Since most resonances of the remainder of the protein are visible in the spectra of the ligand-bound state, we concluded that the study of the dynamic properties of this complex would provide additional insight into the behaviour of the L1 loop and residues in the active site. ^{15}N R_2, R_1 and ^1H-{^{15}N}-heteronuclear NOE relaxation parameters of the monomeric *Pa*DDAH enzyme complexed with 2-chloroacetamidine were measured at 600 MHz (see Chapter 2, Section 2.5 for details of the experiments). The results were analysed following the reduced spectral density mapping procedure (Farrow *et al.*, 1995). $J(0)$, $J(\omega_N)$ and $J(0.87\omega_H)$ have been estimated in the same manner as previously described for the apo-state of monomeric *Pa*DDAH and the values are plotted as a function of residue number in Figure 6.15. For comparative purposes, the values of $J(0)$, $J(\omega_N)$ and $J(0.87\omega_H)$ obtained for the apo-protein are also plotted. Clearly, both the ligand-bound and the apo-state of *Pa*DDAH display remarkably similar spectral density values. In the bound state of the protein, the presence of motions on the ms to μs timescale, indicated by $J(0)$ values higher than one standard deviation from the mean are observed for residues Phe 2, Val 62, Gly 116 and for residues 145-146 and 164-166 in the L6 loop (His 162 broadened beyond the detection limit and could not be assigned in the bound state). Incidentally, most regions that experience motions in the ms to μs timescales in the free protein still undergo conformational exchange contributions in the ligand-bound state.

With respect to the higher frequencies *i.e.* $J(0.87\omega_H)$ and $J(\omega_N)$, sensitive to motions occurring at the ps and ns timescale, they are present in loops L1 (17-24), L3 (79 and 81), L6 (158), for residues 180 and 198 which lie in 3_{10}-helices (Figure 6.15b,c). Following the ligation of 2-chloroacetamidine, the regions of the protein experiencing ps to ns timescale motions are the same as those of the free protein. In particular, the presence of a small covalently linked molecule in the active site has not affected the motional regime of the L1 loop over the active site.

The view that emerges from these observations is that although the changes in chemical shifts observed upon 2-chloroacetamidine binding attest to some conformational rearrangements and changes in the chemical environment of several

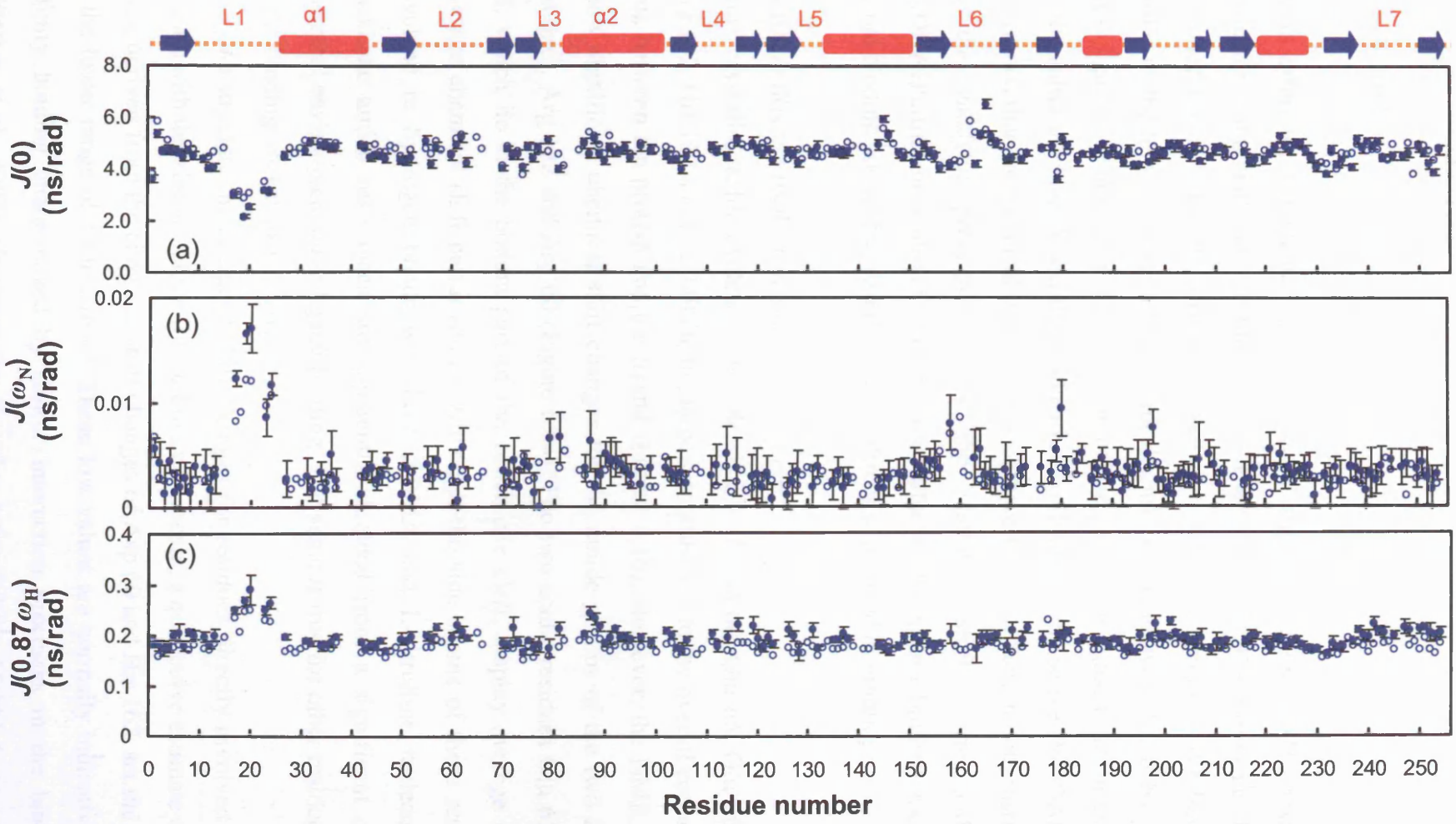

Figure 6.15. Values of the reduced spectral densities $J(\omega)$ of R40E-R98H *Pa*DDAH bound to 2-chloroacetamidine (full circles) and of apo-R40E-R98H *Pa*DDAH (open circles) estimated at (a) $J(0)$, (b) $J(\omega_N)$ and (c) $J(0.87\omega_H)$ plotted as a function of residue number and extracted at 600 MHz. Secondary structure elements are indicated at the top of the Figure.

nuclei, the timescales of motional regimes are preserved in most parts of the protein compared to the free protein.

6.5 Discussion

In an effort to characterise the ligand-bound state of the *Pa*DDAH enzyme, small molecule compounds designed to inhibit the mammalian DDAH enzymes have been screened for their ability to bind the bacterial DDAH homologue. We have applied NMR methodology to study the chemical shift perturbations induced by several of these ligands in order to map the protein binding site. The initial titration experiments of the substrate ADMA and the product L-citrulline into the active *Pa*DDAH enzyme confirmed that the chemical shift changes observed were in fact, in both instances, due to the catalytic product. Namely the chemical shifts correspond to the *Pa*DDAH:citrulline complex. The regions affected by ligand binding included the actual binding pocket but also regions not directly involved in binding.

Residues directly implicated in ligand contacts

From crystallographic evidence, the side chains of Asp 60, Glu 65, Glu 66, Arg 85, Arg 132, His 162 and Cys 249 are the largest contributors to the overall contact surface area between the protein and the ligand (Figure 6.16). However, the NMR data only show significant chemical shift changes for the amide groups of the two active site residues, Arg 132 and Asp 60 (Figure 6.3a). The two acidic residues Glu 65 and Glu 66, which lie at the bottom part of the active site cleft, display average to below-average chemical shift perturbations. Although the side chains of these residues are involved in hydrogen bonds with both ADMA and L-citrulline molecules, their backbone amide resonances are apparently shielded from a significant change in chemical environment upon ligand binding. The same is true for other residues located in the binding pocket such as Arg 85.

In trying to relate the chemical shift changes for residues directly involved in ligand contact with the kinetics of the ligand binding process, a qualitative estimate of the off-rates, derived from the chemical shift changes of Asp 60 and His 162, set the k_{off} value at the lower range of 44 to 276 s^{-1}. These low values are normally indicative of high affinity binding characterised by slower interaction processes in the binding site (Günther *et al.*, 2002). However, the K_d value estimated for ADMA from a fit of the

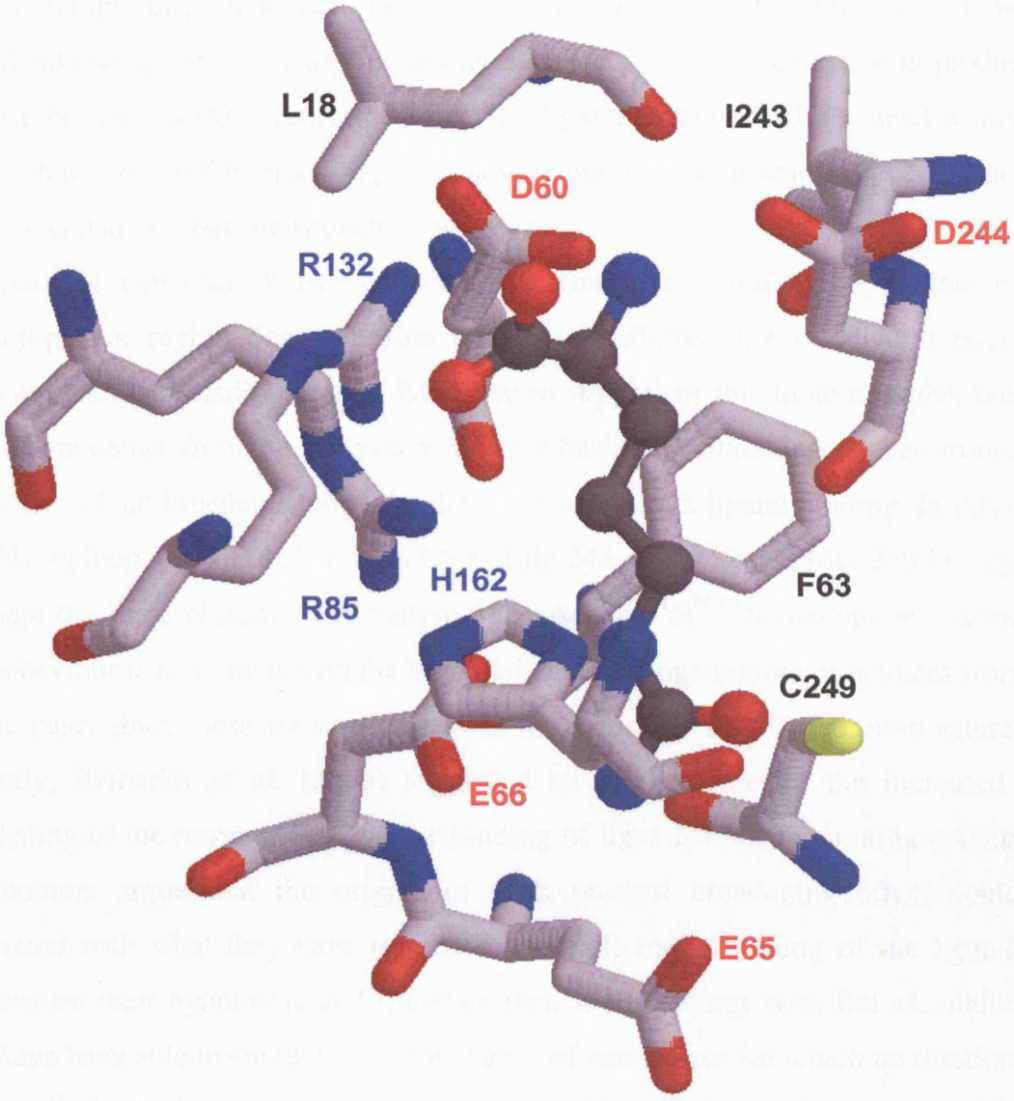

Figure 6.16. 3D view of the active site region of *Pa*DDAH complexed with L-citrulline (pdb code: 1h70). Residues involved in ligand-contact are shown in CPK colouring. Note, the O atom of Ser 249 in the *Pa*DDAH X-ray structure has been coloured yellow to reflect the presence of the S atom of Cys 249.

ligand-induced chemical shift changes as a function of ligand-to-protein ratio is only in the high micromolar range, which is not characteristic of a tight-binding ligand. However, Equation 6.3 states that $K_d = k_{off}/k_{on}$. For a low k_{off} rate with a K_d value around 10^{-3} M, the k_{on} rate must be around 10^6 L mol^{-1} s^{-1}, a value which is three orders of magnitude lower than the diffusion-controlled rates. A low k_{on} rate could reflect a low availability of protein conformers able to bind the ligand with high affinity, which would ultimately result in a poor interaction. In a similar manner, non-productive binding between several conformers and the ligand would entail the need to reverse these interactions before a subsequent relevant binding event could proceed. Such an event would also lower the apparent on-rate.

The possible existence of two or more conformational states of the protein in the binding-pocket region does not rule out the possibility that the ligand could be sampling alternate binding modes. Evidence in support of this more complex binding mechanism comes from the analysis of specific backbone amide resonances around the active site which broaden beyond the detection limit upon ligand binding. In this case, Gly 116 in loop L4, Thr 165 in loop L6 and Ile 243, Val 247 and Met 250 in loop L7, all share the same characteristic pattern of broadened $^1H^N$,^{15}N correlations. Although this behaviour is consistent with the intermediate exchange regime, it deviates from the classic case, since these resonances do not sharpen as a result of protein saturation. Recently, Reibarkh *et al.* (2006) suggested an explanation for the increased line broadening of the resonances upon the binding of ligands under saturating conditions. The authors argue that the origins of such marked broadening effect could be associated with what they have termed as multiple-mode binding of the ligand. To substantiate their hypothesis and illustrate their own findings with Bcl-xL inhibitors, they have been able to simulate the line shapes of resonances for which no titration end point could be found in the presence of excess of ligand. In our experiments, ADMA and L-citrulline are clearly in large excess with respect to the protein concentration. Following the interpretation of Reibarkh *et al.* (2006), the complete broadening of the set of NH resonances mentioned above indicates that ADMA and L-citrulline might adopt two or multiple bound-states of comparable energy. This mechanism of binding together with the possible low availability of conformers of *Pa*DDAH, which are competent to bind ADMA with high affinity, provide an explanation for the high micromolar K_M and K_d values of ADMA estimated for this enzyme *in vitro*. This view is compatible with the notion that the free enzyme samples multiple conformations

until those resembling the transition state are achieved, allowing catalysis to proceed (Tousignant and Pelletier, 2004). In this regard, it is noteworthy that residues in the active site pocket that completely broadened in the 2D-HSQC spectra upon ligand binding are also residues identified as undergoing chemical exchange in the apo-form of the enzyme as shown by the ^{15}N relaxation studies performed.

The interpretation of the dynamic properties of *Pa*DDAH could also be extended to explain the behaviour of several human DDAH I substrate-based inhibitors we have titrated into the bacterial *Pa*DDAH. Upon SR 140, SR 227 and SR 257 binding to *Pa*DDAH, the number of observed resonances that undergo extreme broadening around the active site region increased when compared to the number that "disappear" during ADMA or L-citrulline titrations. It is tempting to correlate this fact with the inability of these substrate-based inhibitors to bind *Pa*DDAH with high affinity.

However, the specificity of these inhibitors for the mammalian enzymes rather than the bacterial enzyme has an altogether different basis. The ongoing research led by collaborators at Birkbeck College aiming to solve the X-ray structure of the human DDAH I has very recently provided us with an early view of the crystal structure of the human DDAH I. It reveals a larger active site cavity in the human enzyme in comparison to the bacterial homologue, which would enable the binding of "longer" substrate-based inhibitors, such as SR 227 and SR 257. This difference would ultimately explain the much lower IC_{50} of these compounds derived against hDDAH I (Rossiter *et al.*, 2005).

Long range conformational changes upon ligand binding

The titration of ADMA and L-citrulline to the active site of *Pa*DDAH induced some long range conformational adjustments made evident as various regions of the protein not directly involved in the binding event displayed significant chemical shift changes. These regions included the L1 loop residues (17-25) over the active site as shown by the rather large chemical shift deviations of residues Gly 17 and Gly 24 (Figure 6.3). Evidence from the *Pa*DDAH X-ray structure suggests that only one hydrogen bond contact is made between the carbonyl group of Leu 18 and the amide group of L-citrulline or ADMA. However Leu 18 and adjacent residues experience severe line broadening of their amide resonances beyond the detection limit suggesting that this region undergoes relatively large amplitude exchange processes. A qualitative estimate of the off-rates for Gly 17 and Gly 24 range from 600 to 860 s^{-1}, values which are

higher than the ones derived for residues directly involved in ligand binding such as His 162 and Asp 60. The k_{off} values are a rough estimate and current work is aiming to simulate the line shapes of the titration in order to obtain a better view of the kinetics of ADMA binding.

However, the existence of conformational exchange occurring for residues in the L1 loop contrasts with the static picture provided by the crystal structure of *Pa*DDAH whereby a well resolved L1 loop over the active site locks into position upon ligand binding. This view appears unrealistic when challenged by the data we have gathered by NMR. The temperature crystallographic B-factors measured for this region would in principle agree with the notion that the L1 loop dynamics are quenched upon the binding of either ADMA or L-citrulline. However, even upon the binding of a covalently linked ligand such as CyNO the loop still experiences conformational exchange uncoupled from the off-rates of ligand binding. It was interesting to note that the covalent binding of the smaller compound 2-chloroacetamidine to the reactive Cys 249 did not affect the dynamics of the loop which still samples ps to ns timescale motions. In this context, the larger size of the ligand accommodated in the active site region is in fact likely to be responsible for causing the subsequent broadening of the $^{1}H^{N}$,^{15}N correlations corresponding to residues in the L1 loop. In a similar way, residues in helices α1 and α2 for which backbone resonances display above average chemical shift changes when CyNO, ADMA or L-citrulline are titrated, are not perturbed when 2-chloroacetamidine is bound.

Taken together these results provide a different view of *Pa*DDAH ligand-bound state in which the L1 loop believed to close down upon the active site, in reality samples two- or multiple conformations in the intermediate exchange regime. The biological implications underlying the presence of conformational exchange in the ligand-bound form of the enzyme are still unclear and it is difficult to speculate at this stage prior to a more detailed view of the dynamics of the complex. Quantifying the exchange rates of the conformations sampled by this loop will provide a better understanding as to how this movement relates to the binding rates and catalytic rates. To initiate this analysis, the ^{15}N relaxation studies performed on the 2-chloroacetamidine complex showed that motions in the millisecond to microsecond timescales are still present in the L6 loop where the active site residue His 162 lies. These results are reconciled with other studies on enzyme dynamics, where apo- and bound states of the protein display

a conserved network of backbone motions in the μs to ms timescale (Beach *et al.*, 2005). More interesting still are the possibilities of linking CPMG relaxation dispersion experiments of the enzyme in the presence of a low stoichiometric amounts of ligand, in order to avoid the complete broadening of the informative resonances, with the simulation of the line shapes obtained during the titration experiments, which might reveal the degree of complexity of the binding mechanism such as has been uncovered in other studies (Mittag *et al.*, 2004; Reibarkh *et al.*, 2006).

The following and final chapter will further discuss these possibilities and will present an overall view of the results described in this thesis.

Chapter 7

Conclusions and perspectives

Abstract

This chapter provides an overview of the main results described in this work, which aimed to characterise the bacterial 58 kD *Pa*DDAH enzyme through the application of NMR methodologies. Due to the intrinsic physico-chemical properties of the native homodimeric enzyme, the initial goals of the project had to be redirected towards the engineering of a monomeric protein which proved more amenable for NMR studies. The findings presented here provide a first insight into the dynamical properties of the guanidino-modifying class of enzymes. In particular, the active site loop L1 displays ps to ns timescale motions in the apo-state of *Pa*DDAH, whilst motions in the μs-ms timescale occur in the active site region. Furthermore, the titrations of *Pa*DDAH with several ligands have demonstrated that a subset of amide resonances for residues located in the L1 loop and active site regions undergo severe line broadening upon ligand binding. These results have been interpreted to imply that the bound-state of *Pa*DDAH is subject to conformational fluctuations even upon saturation of the enzyme. Finally, the suggestion for several complementary studies to help in the understanding of the DDAH enzyme's structure-function characteristics will be discussed.

7.1 Pharmacological interest of the DDAH enzymes

The asymmetric methylarginines, ADMA and L-NMMA, are important endogenous inhibitors of all three isoforms of nitric oxide synthases (NOS; Leiper and Vallance, 1999). By competitively displacing L-arginine from the substrate binding site of NOS, ADMA interferes with many of the physiological roles of nitric oxide (NO), impairing the overall vascular homeostasis. The significance of ADMA accumulation in plasma and tissues has long been recognised in humans, as it has been linked to several cardiovascular disease states, including renal failure, diabetes and pulmonary hypertension (Vallance and Leiper, 2004). In fact, high levels of ADMA have now been proven to be a strong indicator of premature cardiovascular risk (Böger, 2004). The enzymes DDAH I and DDAH II are responsible for metabolising asymmetric methylarginines. By controlling the local cellular levels of ADMA and L-NMMA, DDAH activity is potentially involved in the regulation of NO production, although little evidence exists to implicate the dysfunction of these enzymes as the direct cause for the accumulation of ADMA *in vivo.*

In order to address the deleterious effects of the impairment of DDAH I activity *in vivo*, Vallance and co-workers have recently been able to produce transgenic DDAH I +/- knockout mice (Leiper *et al.*, *submitted*) and novel small molecule inhibitors of DDAH I (Rossiter *et al.*, 2005). Their results demonstrate that the loss of DDAH I function, through either genetic or chemical biology approaches, leads to the accumulation of ADMA and the reduction of NO signalling. Ultimately, the pathophysiological effects observed in the examined mice include endothelial dysfunction and pulmonary hypertension, which are clear signs of acute ADMA accumulation. These findings are the first compelling evidence linking the elevated levels of ADMA commonly found in disease states to DDAH I dysfunction *in vivo* (Leiper *et al.*, *submitted*).

It is noteworthy that DDAH activity should be intimately involved in the regulation of NO signalling. A recent report has suggested that a mechanism of negative feedback existed, in which DDAH could be inhibited *in vitro* by the nitric oxide S-nitrosylation of the active site Cys 249 (Leiper *et al.*, 2002). The authors conclude that under conditions of elevated NO in cells, DDAH inhibition could be a potential mechanism to up-regulate the ADMA levels and consequently limit further production of NO which is in some cases detrimental. The extent of the "cross talk" between the

ADMA:NOS:DDAH pathways is currently the target of on-going research (Tran *et al.*, 2003). Future work in this field is aiming towards a better understanding of the signalling processes involved in the synthesis and the regulation of methylarginine turnover and how these are ultimately integrated (Leiper *et al.*, *submitted*).

7.2 The engineering of a monomeric enzyme

The pharmacological interest posed by the study of DDAH enzymes has prompted a collaborative effort between UCL and Birkbeck College towards the structural characterisation of these enzymes. The fact that neither recombinant mammalian DDAH I and DDAH II enzymes could be expressed in sufficient amounts for structural studies redirected the focus of our research on the bacterial DDAH homologue from *Pseudomonas aeruginosa*. This enzyme was found to be part of a class of guanidino-modifying enzymes that share a conserved β/α-propeller fold (Shirai *et al.*, 2001). More recently, an interesting comparative study between three bacterial hydrolases from this superfamily, including *Pa*ADI and *Pa*DDAH, offered an insight into their common origins (Lu *et al.*, 2006). The authors suggest that gene recruitment in *P. aeruginosa* might have been involved in the generation of a secondary metabolic pathway to facilitate environmental adaptation (Lu *et al.*, 2006). Due to their presence in bacterial strains, these enzymes have also been suggested to be potential targets for new classes of anti-bacterial agents (Vallance *et al.*, 2005), which further validates the structural studies of the *Pa*DDAH homologue.

From an NMR perspective, the 58 kD *Pa*DDAH enzyme provided the laboratory with an ideal case-study for the application of NMR methodology to "larger" protein systems. Furthermore, this project was in turn an opportunity to access the binding schemes of several DDAH ligands that were provided by collaborators at the Clinical Pharmacology Department, to derive possible class-specific inhibitors of mammalian and bacterial DDAH isoforms and ultimately, to probe the dynamical properties of the free and ligand-bound enzyme. However, in the course of our investigations of the 58 kD *Pa*DDAH enzyme, the intrinsic physico-chemical properties of the homodimer precluded the sequence-specific resonance assignments, which are a primary requirement for any structural characterisation by NMR. Although the application of TROSY-based pulse sequences in conjunction with the perdeuteration of the

recombinant protein did improve the quality of the spectra, it was still not possible to obtain extensive unambiguous assignments for the native enzyme due to the lack of a complete tally of expected NH cross peaks (Plevin, 2003). A possible explanation for such a drawback included the intrinsic monomer-dimer equilibrium dynamics at the interface region, which could account for the broadening of the resonances of nuclei located in that region of the structure. Such conformational exchange, which has been shown to be reminiscent of alternative protomer packing configurations in other systems (Lukin *et al.* 2003; Yung *et al.* 2003), led us to pursue a monomerisation strategy of the enzyme. This approach was founded on the careful analysis of the interface residues and the hypothesis that the dimerisation of *Pa*DDAH was not a requirement for its biological function, but rather an artefact of the high concentration of protein needed for NMR and in crystallisation trials. The mutation of two residues in this interfacial domain, which were intended to disrupt two reciprocal salt-bridges, proved successful in generating an active R40E-R98H *Pa*DDAH monomeric enzyme (Plevin *et al.*, 2004). The enzymatic assays performed of the double mutant *Pa*DDAH yielded a catalytic rate constant k_{cat} similar to the WT protein ($k_{cat} \sim 0.5\ s^{-1}$), suggesting that the enzymatic activity of the R40E-R98H *Pa*DDAH has not been compromised by the two point mutations (see Chapter 6, Section 6.2). Furthermore, the similarity between the preliminary [^{1}H,^{15}N]-HSQC spectra acquired of the WT and mutant proteins showed that no substantial disruption of the overall fold of the enzyme has occurred (Figure 3.8). Subsequently, the improved quality of all heteronuclear 2D and 3D spectra recorded for the monomeric protein allowed for the complete backbone resonance assignments of the enzyme (Magalhães *et al.*, 2004). With the assignments of the monomeric protein in hand, we were able to assign tentatively the amide cross peaks for the WT *Pa*DDAH. Interestingly, the absent cross peaks in the WT spectra mapped to parts of the extended β-sheet across the subunits, to residues around the active site catalytic triad and to residues in the L1 loop (see Chapter 4, Section 4.4, Figure 4.11). This finding confirms our initial surmise that chemical exchange does occur at the interface region of the homodimer, suggesting that the monomer-dimer equilibrium could be at the basis of conformational rearrangements on the intermediate exchange rate regime which lead to the broadening of the amide resonances.

7.3 Considerations on enzyme dynamics and function

To date most of the information available on the kinetics and the thermodynamic properties of an enzyme-catalysed reaction comes from extensive studies of substrate to product turnover. However, it is clear that enzyme function depends not only upon the transition state energy of the reactants but also upon the kinetic and the thermodynamic energy landscape of the conformational rearrangements of the enzyme along the reaction pathway. In this respect, ^{15}N relaxation studies performed of *Pa*DDAH are a starting point in the characterisation of the internal dynamics of this class of enzymes.

The current investigation has revealed that the secondary structure elements of *Pa*DDAH display a high average value of NH bond vector generalised order parameter (S^2 ~ 0.87) attesting to the overall rigidity of the β/α-propeller fold. Despite this fact, four loops located around the entrance of the active site have been shown to sample motions in the ps to ns timescale (see Chapter 5, Sections 5.3 and 5.5; Figures 5.8 and 5.13). In particular the NH bond vectors corresponding to residues in the L1 loop located directly over the active site (residues 14-25) display the lowest S^2 values (S^2 ~ 0.6). Moreover, slower motions in the μs to ms timescale have also been identified in the *Pa*DDAH active site region. More specifically, the model-free analysis together with the field-dependent linear regression analysis have derived R_{ex} terms for residues Ser 61, Val 62 and Glu 65 located in loop L2, for Gly 116 in loop L4, for the active site residue His 162 and the surrounding residues Lys 164 and Thr 165, and for Ser 248, Cys 249 and Ser 251 located in loop L7 (see Chapter 5, Sections 5.3 and 5.4). Interestingly, although the sequence similarity scores between mammalian DDAH enzymes and the bacterial homologue are rather low (~20-40%), the inspection of the sequence alignment between rat DDAH I, human DDAH I and *Pa*DDAH proteins reveals that residues situated in the binding pocket of the enzyme that have been found to undergo exchange contributions to their relaxation rates are in fact remarkably conserved in these three enzymes. The only exception is residue Ser 251 in *Pa*DDAH, for which the counterpart in the mammalian enzymes is a threonine residue.

The importance of the conformational exchange occurring at slower timescales has been further underlined in other enzymes as being fundamental to promote catalysis. In the NMR studies of dihydrofolate reductase (DHFR), cyclophilin A, and RNase A (Osborne *et al.*, 2001; Eisenmesser *et al.*, 2005; Beach *et al.*, 2006), the conformational

exchange rates k_{ex} derived from the apo-state of the protein have been quantitatively linked to the rate-limiting step of the catalytic mechanism, k_{cat}. Eisenmesser and colleagues further discuss the fact that protein dynamics associated with catalysis is a "built-in property" of the enzyme that is also manifested in the free protein (Eisenmesser *et al.*, 2005). However, in the case of *Pa*DDAH, without a better estimate of the exchange rates for the residues around the active site, one can not affirm that the conformational rearrangements around the binding-pocket are rate-limiting for the overall function of the DDAH enzymes. In fact the rather low k_{cat} values of the DDAH enzymes (Stone *et al.*, 2005a) around ~0.5 s^{-1} would predict that only very slow conformational changes of the enzyme are prone to be the determinant step in the catalytic reaction.

The titration of several ligands into *Pa*DDAH aimed to characterise the ligand-bound state of the protein. Our findings show that upon saturation of the enzyme with either ADMA or L-citrulline, a subset of the amide resonances for residues located in the L1 loop were subject to severe line broadening resulting in ablation of cross peak intensity (see Chapter 6, Section 6.2). These observations suggest that the L1 loop undergoes a change in dynamic motional regime induced upon ligand binding. This situation has been shown to occur in other systems (Eisenmesser *et al.*, 2005). For example, in the case of cyclophilin A, a loop (comprising residues 65-84) in the free enzyme displays faster exchange rates (k_{ex} ~ 2300 s^{-1}) with respect to the overall catalytic rates (k_{cat} ~ 1100 s^{-1}). However upon substrate binding the motional regime of this loop is altered and it fluctuates collectively with the rest of the protein as catalysis occurs (Eisenmesser *et al.*, 2005).

Furthermore, the ligand binding studies of *Pa*DDAH have also shown that several amide resonances corresponding to residues directly involved in interactions with the ligand exhibit severe line broadening with increasing ligand concentration to the point where several NH resonances remained undetected, even upon saturation of the enzyme (see Chapter 6, Section 6.2). It is interesting to note that these resonances around the active site have also been identified as undergoing chemical exchange in the apo-state of *Pa*DDAH. This observation is further supported by the ^{15}N relaxation experiments performed on 2-chloroacetamidine-bound *Pa*DDAH, in that both free and ligand-bound states of the protein display a conserved network of backbone motions in

the μs to ms timescale for loop L6, where the active site residue His 162 lies (see Chapter 6, Section 6.4).

Despite our best efforts, we were unable to obtain a ligand-bound conformation of *Pa*DDAH, for which all the NH resonances for the L1 loop and the active site region could be observed under saturating conditions. This situation deterred us from pursuing further ^{15}N relaxation experiments of the bound-state of the enzyme due to our inability to fully interpret the behaviour of these resonances. Recently, there has been evidence suggesting that the ablated NH resonances under saturating conditions could be induced by two- or multiple-mode binding of the ligand in the active site (Reibarkh *et al.*, 2006). The possible existence of multiple ligand bound conformations in the case of *Pa*DDAH could help explain the high micromolar affinity for ADMA and the low catalytic efficiency of this enzyme, although these properties are not necessarily causually related.

7.4 CPMG relaxation dispersion and line shape analysis

The combination of ^{15}N relaxation studies and the analysis of chemical shift changes in the mapping of *Pa*DDAH binding site have given us an initial view of the intrinsic dynamics of *Pa*DDAH. To build on these preliminary studies, it would be ideal to quantify the conformational exchange rates of *Pa*DDAH and to have a more reliable estimation of the ligand binding on- and off-rates. These values of k_{ex}, k_{on} and k_{off} rate constants would be essential in order to compare them with the catalytic rate of the enzyme k_{cat}. To achieve this comparison, an interesting prospect is the application of CPMG relaxation dispersion experiments to probe the motions of *Pa*DDAH on the μs and ms timescale upon ligand binding. By choosing to study the dynamics of the enzyme induced by low stoichiometric amounts of ligand, information on the exchange rates undergone by the L1 loop could be obtained, as the amide resonances for the residues in this region would not have completely broadened as in the saturated state. Under these circumstances, the exchange rate k_{ex} is related to the k_{on} and k_{off} rates according to $k_{ex} = k_{on}[L] + k_{off}$, where [L] is the concentration of ligand. Once k_{ex} is obtained then the ligand binding rate constants can be obtained from the equilibrium dissociation constant, K_d ($K_d = k_{on}/k_{off}$; Rule and Hitchens, 2006).

However, the separation of protein internal dynamics and dynamics associated with ligand binding might also be achieved through the independent analysis of the ligand NMR titration experiments. The simulation of line shapes can provide good estimates of the k_{off} rates for the binding kinetics and report on the complexity of the binding mechanism (Günther and Schaffhausen, 2002).

The complementarity of these two types of analysis has been successfully demonstrated in a series of studies of the N-terminal src homology 2 domain (SH2) of phosphatidylinositide 3-kinase (Mittag *et al.*, 2003; Mittag *et al.*, 2004). CPMG relaxation dispersion measurements revealed the existence of low populated conformers that could not be detected by simulating the line shapes (Mittag *et al.*, 2003). Moreover, the observed line shapes were fit to different kinetic binding models consistent with the existence of structural intermediates in the binding pathway and providing a basis for understanding the different binding affinities for different ligands (Mittag *et al.*, 2004).

To complement the dynamic studies of the *Pa*DDAH enzyme reported here, further experiments, such as those mentioned above, are currently part of ongoing collaborations with Dr T. Frenkiel at the National Institutes of Medical Research in London and Dr U. Günther at the University of Birmingham.

7.4 An integrated view from X-ray crystallography, NMR spectroscopy and molecular dynamics (MD)

The *Pa*DDAH crystal structure offered us a first insight into the structural intricacies of the class of guanidino-modifying enzymes (Murray-Rust *et al.*, 2001). Comparison with other X-ray structures of this class of proteins reveals the basis of the specificity of these enzymes for a specific substrate (see Chapter 1, Section 1.8; Galkin *et al.*, 2004; Lu *et al.*, 2006). In this sense, the data acquired from X-ray crystallography is invaluable, as it can provide different and altogether informative snapshots of the systems under consideration. In particular, when several ligands are available, as is the case with DDAH, it would be extremely interesting to obtain additional structural information on how these ligands are accommodated in the active site. Unfortunately, not all systems are investigated in such a manner and this caveat undermines the possibility for exploring the information contained in the structural heterogeneities that could arise from the collection of a crystallographic ensemble. This point has been

elegantly argued by Furnham and co-workers when discussing the need for a more suitable representation of the macromolecular crystal structure, whereby the X-ray structures deposited in the PDB data bank should be accompanied by a range of structural models (Furnham *et al.*, 2006).

In adding to the information provided by the *Pa*DDAH crystal structure, the complexity of the binding mechanism of *Pa*DDAH would certainly benefit from a more detailed investigation of the side chain dynamics implicated in interactions with the ligand. For this particular study, recent developments in the NMR field and in the isotopic labelling strategies have improved the sensitivity of deuterium-based spin relaxation studies of selectively ^{2}H-labelled methyl groups. Several research groups have now contributed to this active area of research addressing the role of side chain dynamics in molecular recognition, protein folding and stability (for reviews see Tugarinov and Kay, 2005; Kay, 2005); however currently no particular study has focused on enzyme side chain dynamics.

To aid the investigation of the dynamical properties of *Pa*DDAH, we have also considered the possibility of applying molecular dynamics (MD) to understand the role of the active site loop dynamics upon ligand binding. Classic molecular dynamic simulations for up to 10 ns have revealed the existence of correlated motions in the backbone of DHFR (Radkiewicz and Brooks, 2000). In particular, coupled motions that appear in the reactive complex disappear in the product complexes, indicating that these motions may have a role in catalysis. In the studies of the HIV-1 protease, an all-atom simulation revealed that the preferred structure, either closed, semiopen, or fully open forms of the protein, was strongly dependent upon the presence of an inhibitor (Hornak *et al.*, 2006). Several other examples of MD simulations that investigate the dynamics of proteins upon ligand binding exist in the literature and it would be of interest to apply this procedure to *Pa*DDAH.

On another note, so far our studies have not addressed the thermodynamics of ligand binding. In this sense, molecular dynamic simulations are also likely to be of increasing importance for the derivation of per-residue entropies for side chain residues (Homans, 2005). Ultimately, these studies could enable the determination of the factors governing affinity which could help explain the high K_M values of

*Pa*DDAH for its substrates and the low affinity displayed by the DDAH I inhibitors towards the bacterial homologue.

In summary, it is clear that apart from the ongoing collaboration efforts directed towards the determination of a more quantitative measure of the internal dynamics of *Pa*DDAH, the integrated approach addressed here would certainly provide a more complete inspection of the structure-function characteristics of the enzyme. Moreover, it is our belief that an even more challenging and certainly very interesting avenue for the understanding of these enzymes would be the "taming" of the human DDAH I enzyme for the application of NMR methodology. In this vein, our colleagues at Birkbeck College have recently solved the X-ray structure of DDAH I with a very limited amount of protein (Leiper *et al.*, *submitted*). Any comparative structural study between bacterial and mammalian homologue enzymes will ultimately add to the understanding of how the subtle differences in the arrangement of active site residues modulate the different ligand binding affinities.

Appendix

Table A.I. Details of the datasets used in each fitting procedure to a monomer-dimer equilibrium model

*Pa*DDAH isoform	Molecular weight (Da)	Fittings	Concentration (mg/ml)	Rotor speeds (10^3 rpm)	Ka_2[a]	K_d (M)[b]
R98H	29198	1	0.92	15, 20, 24	0.218	9.49×10^{-5}
			0.61			
			0.32			
		2	0.92	11, 15, 20	0.262	7.90×10^{-5}
			0.61			
			0.32			
		3	1.62	11, 15, 20	0.205	10.1×10^{-5}
			0.92			
			0.32			
		4	2.96	11, 15, 20	0.281	7.36×10^{-5}
			1.62			
			0.82			
R40E	29190	1	1.46	11, 15, 20	0.276	7.5×10^{-5}
			0.71			
			0.34			
		2	1.46	15, 20, 24	0.248	8.35×10^{-5}
			0.71			
			0.34			
		3	1.46	11, 15, 20	0.356	5.81×10^{-5}
			0.61			
			0.34			
		4	1.46	15, 20, 24	0.274	7.55×10^{-5}
			0.61			
			0.32			
Q43R	29245	1	0.17	11, 15, 20	0.517	3.99×10^{-5}
			0.96			
			1.68			
		2	0.17	15, 20, 24	0.397	5.20×10^{-5}
			0.56			
			1.68			
		3	0.56	11, 15, 20	0.485	4.26×10^{-5}
			0.96			
			1.68			
		4	0.56	15, 20, 24	0.425	4.86×10^{-5}
			0.96			
			1.68			
		5	0.17	15, 20, 24	0.471	4.37×10^{-5}
			0.96			
			1.68			
		6	0.37	15, 20, 24	0.459	4.50×10^{-5}
			0.96			
			1.68			

[a]A dataset is composed of the sedimentation equilibrium interference curves obtained at three different concentrations recorded at three different rotor speeds. These nine sedimentation equilibrium curves were fit simultaneously to the monomer-dimer equilibrium association model to obtain Ka_2 values. [b]K_d values were derived using Equation 2.3 (see Chapter 2, Section 2.4)

Table A.II. Chemical shifts of ^{1}H, ^{15}N, $^{13}C_{\alpha}$, $^{13}C_{\beta}$ and ^{13}CO resonances of R40E-R98H *Pa*DDAH corrected for the deuterium effect (ppm)

Residue		N	HN	C_{α}	CO	C_{β}	Residue		N	HN	C_{α}	CO	C_{β}
1	Met	121.054	7.576	56.391	176.229	36.202	51	Leu	129.26	8.833	53.973	176.053	43.802
2	Phe	119.5	9.51	56.379	175.779	40.472	52	Leu	130.3	9.442	53.43	-	41.32
3	Lys	120.089	9.566	57.65	176.425	36.027	53	Pro	-	-	63.152	177.861	32.479
4	His	119.196	8.709	55.479	172.175	33.572	54	Asp	119.822	8.092	54.097	176.198	44.188
5	Ile	116.079	8.021	57.245	177.381	40.58	55	Glu	119.822	8.092	54.097	176.198	44.188
6	Ile	124.498	8.559	60.892	174.077	41.458	56	Arg	121.382	8.649	58.447	178.958	30.437
7	Ala	128.4	8.914	51.653	174.095	24.415	57	Phe	116.243	8.178	58.256	175.957	30.968
8	Arg	121.874	9.799	56.603	172.103	34.254	58	Pro	113.922	7.506	56.28	-	41.665
9	Thr	123.875	7.882	62.768	-	72.02	59	Asp	-	-	66.864	177.822	33.627
10	Pro	-	-	62.748	174.84	32.896	60	Ser	113.826	8.574	55.478	177.535	43.072
11	Ala	126.06	8.15	51.474	177.979	20.195	61	Val	112.908	8.172	62.083	175.769	64.093
12	Arg	124.942	10.35	60.433	180.147	28.851	62	Phe	121.033	7.141	63.406	176.119	31.741
13	Ser	113.236	8.919	59.969	175.764	63.895	63	Val	125.996	8.262	54.816	177.73	35.289
14	Leu	124.951	7.826	59.547	177.978	43.402	64	Glu	115.857	7.494	63.456	173.781	31.99
15	Val	108.337	7.053	63.77	176.201	31.946	65	Asp	116.378	7.626	59.293	175.833	31.818
16	Asp	120.179	7.695	55.722	176.999	41.927	67	Pro	108.831	6.857	60.531	-	40.638
17	Gly	107.989	7.941	46.478	174.062	-	68	Val	-	-	65.998	173.838	33.554
18	Leu	122.164	8.707	56.589	177.953	42.57	69	Leu	114.303	6.941	61.186	170.995	36.004
19	Thr	114.259	8.215	62.298	180.213	70.784	70	Cys	122.976	8.666	53.878	175.377	45.901
20	Ser	118.173	8.59	59.22	-	65.143	71	Thr	115.834	8.158	61.244	175.481	73.016
21	Ser	-	-	-	-	-	72	Ser	116.608	10.07	61.408	176.337	63.901
22	His	-	-	58.266	176.084	30.59	73	Arg	117.566	7.945	56.777	175.166	34.081
23	Leu	120.227	7.992	56.053	177.559	43.048	74	Cys	109.276	7.115	57.484	171.265	29.573
24	Gly	106.983	7.76	45.037	174.363	-	75	Ala	123.072	8.206	50.628	175.243	21.74
25	Lys	119.106	8.326	54.948	-	33.473	76	Ile	119.942	9.108	60.873	175.222	41.185
26	Pro	-	-	63.579	174.091	32.469	77	Ile	128.817	9.07	61.056	176.73	36.161
27	Asp	123.82	8.928	53.429	176.309	44.008	78	Thr	115.834	8.158	61.244	175.481	73.016
28	Tyr	127.176	9.000	63.352	175.432	39.01	79	Arg	112.397	7.432	55.903	-	30.653
29	Ala	120.213	8.209	56.411	181.862	18.209	80	Pro	-	-	64.914	175.077	32.798
30	Lys	119.142	8.153	59.41	178.43	32.598	81	Gly	111.296	9.058	47.898	174.979	-
31	Ala	123.085	8.804	55.825	183.048	17.013	82	Ala	122.786	7.878	52.327	178.432	18.499
32	Leu	125.613	8.865	59.246	178.556	41.641	83	Glu	126.94	9.018	61.358	179.211	29.81
33	Glu	120.451	7.634	60.632	180.578	30.011	84	Ser	112.516	8.621	61.598	175.375	63.334
34	Gln	119.439	8.725	59.871	179.332	28.318	85	Arg	116.781	7.497	57.684	178.307	30.655
35	His	122.23	8.484	63.974	177.587	31.857	86	Arg	121.568	7.662	61.982	178.874	29.696
36	Asn	116.087	8.629	57.473	178.118	38.701	87	Gly	104.139	8.985	47.264	175.147	-
37	Ala	122.242	8.206	55.953	180.406	18.743	88	Glu	119.106	7.527	59.312	179.699	31.327
38	Tyr	122.618	8.092	62.099	176.048	38.653	89	Thr	107.11	7.522	65.36	176.859	67.842
39	Ile	119.181	8.243	63.287	178.254	35.429	90	Glu	120.522	6.852	59.428	179.554	30.272
40	Glu	118.759	8.265	60.372	179.526	29.786	91	Ile	106.499	6.674	63.57	176.665	37.907
41	Ala	120.32	7.723	55.691	181.64	18.924	92	Ile	121.519	7.201	61.066	175.551	42.216
42	Leu	121.098	7.851	59.078	178.6	42.03	93	Glu	123.635	7.225	61.765	177.305	29.464
43	Gln	116.352	8.275	58.985	177.642	28.969	94	Glu	118.117	8.957	60.928	179.27	29.738
44	Thr	108.382	7.392	63.635	175.359	70.295	95	Thr	116.224	7.547	68.012	176.326	69.088
45	Cys	117.454	7.557	61.012	172.769	28.388	96	Val	119.505	7.659	68.33	178.237	31.998
46	Asp	115.817	8.11	56.891	173.5	39.411	97	Gln	118.241	8.696	59.531	-	28.113
47	Val	105.61	6.595	59.217	174.605	36.18	98	His	115.355	7.311	59.976	176.855	29.76
48	Asp	119.037	7.836	53.402	175.271	43.328	99	Phe	114.302	7.195	61.427	175.391	42.192
49	Ile	120.954	8.575	61.058	176.44	39.573	100	Tyr	119.348	8.01	57.426	-	39.017
50	Thr	127.729	8.841	64.141	171.969	70.41	101	Pro	-	-	65.393	180.197	31.094

Table A.II. *Cont. I*

Residue		N	HN	C_α	CO	C_β	Residue		N	HN	C_α	CO	C_β
102	Gly	113.309	8.926	47.026	175.535	-	152	Gly	104.489	8.538	46.657	171.561	104.489
103	Lys	121.418	8.26	55.279	174.099	34.266	153	Ser	113.229	8.35	58.574	171.688	65.889
104	Val	119.394	7.806	63.061	176.529	33.355	154	Val	119.738	8.512	62.389	177.185	34.367
105	Glu	129.111	8.925	55.187	174.7	32.952	155	Val	128.833	9.204	61.352	174.411	35.465
106	Arg	116.246	8.992	56.263	174.503	33.393	156	Arg	126.904	8.797	57.815	176.111	30.837
107	Ile	120.512	7.597	58.058	175.257	36.474	157	Leu	121.597	8.144	55.319	176.72	44.4
108	Glu	125.841	9.474	55.104	177.096	33.592	158	Glu	120.927	10.70	58.43	178.117	33.589
109	Ala	126.141	9.151	53.131	-	17.62	159	Lys	118.75	9.626	57.967	176.994	34.687
110	Pro	-	-	64.861	177.006	33.421	160	Val	111.021	7.078	60.698	175.517	35.655
111	Gly	113.633	9.044	46.59	172.33	-	161	Leu	120.287	8.13	58.562	175.855	42.947
112	Thr	121.418	8.26	55.279	174.099	34.266	162	His	108.225	7.138	52.968	177.208	33.963
113	Val	119.394	7.806	63.061	176.529	33.355	163	Leu	123.982	10.11	60.399	177.608	44.615
114	Glu	124.291	8.496	55.457	176.667	31.823	164	Lys	107.021	7.182	59.636	176.978	29.969
115	Ala	125.378	9.58	55.868	178.579	17.569	165	Thr	115.616	7.985	67.517	176.646	67.493
116	Gly	110.294	8.418	48.594	173.34	-	166	Gly	104.861	7.375	46.411	169.614	-
117	Asp	117.296	8.145	56.691	174.277	41.475	167	Leu	120.717	6.441	54.516	175.515	49.754
118	Ile	118.929	7.648	61.65	173.11	39.683	168	Ala	118.858	7.972	53.606	175.285	24.324
119	Met	127.527	9.221	54.664	174.057	38.301	169	Tyr	123.366	10.00	58.409	177.11	38.337
120	Met	129.889	9.35	56.291	175.494	35.108	170	Leu	129.525	8.571	53.67	-	41.004
121	Val	132.166	9.263	61.361	175.826	32.434	171	Glu	104.543	4.308	58.867	173.985	28.142
122	Gly	120.545	9.402	48.326	174.323	-	172	His	111.641	8.584	56.696	174.55	26.816
123	Asp	126.969	8.386	54.867	172.883	42.176	173	Asn	108.872	8.113	55.967	173.811	41.435
124	His	119.123	7.87	55.551	172.797	34.619	174	Asn	119.317	7.683	54.586	172.019	40.52
125	Phe	124.901	7.818	56.779	174.496	40.157	175	Leu	128.212	8.76	53.799	175.064	48.128
126	Tyr	122.82	8.832	59.3	175.114	40.121	176	Leu	126.527	9.325	55.298	176.643	42.106
127	Ile	123.186	9.27	60.824	175.611	40.754	177	Ala	120.811	8.058	51.936	174.355	23.938
128	Gly	116.64	10.33	46.495	171.051	-	178	Ala	121.893	8.582	50.229	179.775	24.131
129	Glu	123.374	8.039	56.57	176.16	30.798	179	Gly	108.399	8.851	48.484	177.131	-
130	Ser	126.225	9.201	57.753	174.358	69.59	180	Glu	124.793	9.118	58.213	176.08	30.354
131	Ala	120.016	8.861	54.823	178.898	19.338	181	Phe	116.785	7.363	61.093	176.251	41.016
132	Arg	116.49	8.684	56.134	176.146	31.98	182	Val	114.306	6.908	67.075	176.502	32.474
133	Thr	122.132	7.771	62.642	174.044	70.115	183	Ser	109.183	7.707	58.117	175.555	64.669
134	Asn	120.739	8.168	50.985	175.597	40.32	184	Lys	123.622	7.246	56.01	-	32.169
135	Ala	121.339	8.449	56.352	181.157	18.905	185	Pro	-	-	66.296	178.951	32.465
136	Glu	118.995	8.538	58.628	178.32	28.876	186	Glu	117.702	10.42	60.406	177.582	28.658
137	Gly	110.256	8.842	47.984	175.505	-	187	Phe	115.469	7.302	56.878	176.007	39.48
138	Ala	123.187	8.64	56.626	178.656	18.472	188	Gln	114.997	7.205	59.063	176.411	29.236
139	Arg	117.46	8.221	60.658	180.621	30.237	189	Asp	117.034	8.63	55.833	177.533	41.273
140	Gln	118.789	8.232	60.349	179.291	30.974	190	Phe	118.611	7.602	59.592	175.363	40.588
141	Met	117.674	8.272	57.876	177.445	30.804	191	Asn	118.968	8.871	54.094	173.666	38.123
142	Ile	119.386	8.342	67.365	176.951	37.986	192	Ile	126.436	8.135	62.816	176.546	38.614
143	Ala	120.155	7.635	56.087	181.517	18.233	193	Ile	133.954	9.491	60.697	174.152	37.507
144	Ile	120.026	7.881	66.592	178.079	38.774	194	Glu	126.081	8.302	57.042	176.324	30.306
145	Leu	119.596	8.332	59.49	180.213	40.325	195	Ile	130.072	8.516	56.346	-	37.602
146	Glu	118.156	8.6	60.184	181.573	29.854	196	Pro	-	-	63.404	178.103	32.759
147	Lys	122.517	7.721	59.641	177.945	32.3	197	Glu	123.102	8.873	61.199	179.779	30.289
148	His	115.699	7.328	57.769	174.981	31.927	198	Glu	117.54	9.749	59.738	177.519	29.144
149	Gly	107.637	7.894	47.119	173.896	-	199	Glu	118.461	8.062	55.975	175.388	30.878
150	Leu	121.869	8.145	54.043	174.655	42.471	200	Ser	113.479	7.27	64.19	176.012	64.182
151	Ser	111.879	8.466	58.461	173.841	67.118	201	Tyr	123.322	8.539	60.997	177.543	39.316

Table A.II. *Cont. II*

Residue		N	HN	C_α	CO	C_β
202	Ala	117.776	7.817	53.298	172.984	17.817
203	Ala	115.393	7.022	54.384	173.984	19.183
204	Asn	109	6.524	53.801	174.056	39.652
205	Cys	118.299	6.754	58.638	173.086	33.885
206	Ile	107.369	8.228	59.054	174.235	41.467
207	Trp	124.278	8.687	57.801	177.258	32.227
208	Val	126.784	9.005	62.917	-	34.36
209	Asn	120.9	5.105	55.836	-	-
210	Glu	105.288	5.037	60.202	175.796	26.637
211	Arg	120.173	7.208	56.345	175.633	32.634
212	Val	120.209	8.606	61.411	174.469	34.936
213	Ile	127.956	8.772	60.901	174	38.011
214	Met	126.117	9.025	52.528	-	37.641
215	Pro	-	-	62.058	177.463	32.613
216	Ala	123.532	7.722	52.854	178.074	19.902
217	Gly	108.757	8.41	45.948	174.044	-
218	Tyr	122.39	9.316	57.681	-	39.2
219	Pro	-	-	65.778	180.84	32.972
220	Arg	123.859	10.44	60.885	180.957	29.946
221	Thr	123.41	10.30	69.76	176.221	68.571
222	Arg	121.516	8.388	61.411	178.177	30.232
223	Glu	116.09	7.69	60.033	179.183	29.884
224	Lys	117.775	7.532	60.944	179.226	33.382
225	Ile	117.831	8.025	66.65	179.175	39.146
226	Ala	122.093	8.866	56.078	182.696	18.569
227	Arg	117.543	8.028	59.294	178.023	30.373
228	Leu	117.825	7.427	56.163	177.931	42.888
229	Gly	104.583	7.805	45.936	174.28	-
230	Tyr	116.332	6.453	58.763	175.825	40.55
231	Arg	122.179	7.82	56.998	175.001	30.923
232	Val	126.849	8.434	63.215	176.138	33.348
233	Ile	130.443	9.266	61.817	174.751	41.012
234	Glu	124.931	8.628	55.777	176.987	33.197
235	Val	119.416	8.916	60.289	175.102	34.922
236	Asp	127.139	9.058	55.482	176.623	40.838
237	Thr	114.823	8.305	61.583	178.063	68.622
238	Ser	119.481	8.4	63.469	177.052	63.167
239	Glu	121.659	9.606	59.413	176.74	29.178
240	Tyr	115.568	7.102	61.713	179.496	38.493
241	Arg	124.13	8.279	60.455	179.235	29.996
242	Lys	114.798	6.875	60.395	177.286	33.933
243	Ile	109.875	7.212	61.173	173.979	37.363
244	Asp	112.897	7.521	57.048	174.978	39.895
245	Gly	103.997	8.304	45.917	173.224	-
246	Gly	109.143	6.892	45.067	173.836	-
247	Val	113.247	8.399	66.73	180.345	32.965
248	Ser	122.638	10.43	63.217	176.855	61.388
249	Cys	117.889	7.163	58.842	176.165	30.954
250	Met	109.94	6.571	57.155	175.077	36.866
251	Ser	108.274	7.615	59.756	173.162	67.021
252	Leu	119.603	8.663	55.252	175.723	44.941
253	Arg	122.066	9.405	55.347	173.348	31.48
254	Phe	125.464	8.306	60.397	-	38.683

Table A.III. ^{15}N relaxation parameters of R40E-R98H *Pa*DDAH recorded at 500 MHz

#[a]	R_1[b]	dR_1[c]	R_2[b]	dR_2[c]	NOE	dNOE[d]	#[a]	R_1[b]	dR_1[c]	R_2[b]	dR_2[c]	NOE	dNOE[d]
1	1.244	0.021	12.58	0.35	0.576	0.038	50	1.226	0.036	14.64	0.71	0.715	0.050
2	1.158	0.032	17.87	0.90	0.735	0.066	51	1.211	0.028	13.97	0.53	0.760	0.048
3	1.185	0.050	15.33	1.12	0.831	0.079	52	1.203	0.058	14.56	0.99	0.855	0.086
4	1.238	0.027	16.80	0.50	0.821	0.051	53	-	-	-	-	-	-
5	1.262	0.035	15.77	0.75	0.749	0.054	54	-	-	-	-	-	-
6	1.264	0.045	15.75	0.83	0.773	0.060	55	1.259	0.046	15.87	0.72	0.782	0.053
7	1.202	0.031	14.89	0.56	0.784	0.047	56	1.218	0.041	14.53	0.71	0.891	0.061
8	1.206	0.054	15.81	1.11	0.733	0.061	57	1.235	0.044	15.62	0.82	0.826	0.054
9	1.208	0.040	14.92	0.62	0.849	0.053	58	1.263	0.037	15.11	0.71	0.763	0.045
10	-	-	-	-	-	-	59	-	-	-	-	-	-
11	1.299	0.061	15.54	1.04	0.713	0.065	60	1.252	0.053	14.90	0.90	0.763	0.055
12	1.237	0.041	14.55	0.61	0.725	0.058	61	1.267	0.053	15.93	1.07	0.831	0.061
13	1.258	0.029	15.15	0.51	0.756	0.036	62	1.236	0.049	16.69	1.03	0.766	0.057
14	-	-	-	-	-	-	63	-	-	-	-	-	-
15	1.271	0.027	15.48	0.44	0.802	0.045	64	-	-	-	-	-	-
16	-	-	-	-	-	-	65	1.292	0.039	17.23	0.91	0.829	0.057
17	1.689	0.047	10.62	0.27	0.481	0.025	66	1.260	0.049	16.46	0.97	0.922	0.068
18	1.489	0.025	10.57	0.29	0.492	0.030	67	-	-	-	-	-	-
19	1.567	0.043	9.31	0.42	0.405	0.030	68	-	-	-	-	-	-
20	1.546	0.096	11.51	1.04	0.478	0.058	69	-	-	-	-	-	-
21	-	-	-	-	-	-	70	1.136	0.045	14.12	0.79	0.802	0.077
22	-	-	-	-	-	-	71	1.218	0.030	15.88	0.70	0.734	0.041
23	1.467	0.032	10.72	0.37	0.430	0.031	72	1.275	0.059	15.14	0.98	0.782	0.070
24	1.500	0.038	10.67	0.39	0.472	0.030	73	1.251	0.033	15.21	0.56	0.740	0.047
25	-	-	-	-	-	-	74	1.063	0.023	12.34	0.41	0.618	0.042
26	-	-	-	-	-	-	75	1.262	0.035	15.61	0.57	0.764	0.053
27	1.167	0.032	14.92	0.56	0.803	0.057	76	1.193	0.027	15.99	0.60	0.814	0.058
28	1.306	0.032	15.79	0.54	0.781	0.043	77	1.205	0.043	15.07	0.80	0.806	0.063
29	1.169	0.016	14.96	0.37	0.721	0.033	78	-	-	-	-	-	-
30	1.209	0.018	15.68	0.39	0.758	0.039	79	1.178	0.054	14.67	0.94	0.871	0.090
31	1.212	0.024	16.06	0.46	0.799	0.045	80	-	-	-	-	-	-
32	1.215	0.022	16.31	0.44	0.784	0.044	81	1.155	0.056	13.98	1.03	0.677	0.064
33	-	-	-	-	-	-	82	1.222	0.028	13.60	0.46	0.665	0.036
34	1.184	0.022	16.19	0.48	0.734	0.044	83	-	-	-	-	-	-
35	1.229	0.037	16.81	0.72	0.840	0.056	84	-	-	-	-	-	-
36	1.237	0.023	16.34	0.46	0.770	0.040	85	1.303	0.040	15.13	0.72	0.758	0.047
37	1.205	0.020	16.41	0.46	0.768	0.040	86	1.283	0.034	15.88	0.69	0.807	0.051
38	-	-	-	-	-	-	87	1.462	0.102	16.82	1.93	0.729	0.068
39	-	-	-	-	-	-	88	1.279	0.043	15.57	0.81	0.802	0.058
40	-	-	-	-	-	-	89	1.296	0.034	16.74	0.75	0.755	0.047
41	-	-	-	-	-	-	90	1.166	0.026	15.21	0.43	0.713	0.040
42	1.205	0.026	16.15	0.53	0.747	0.043	91	1.201	0.047	16.07	1.02	0.713	0.071
43	1.231	0.021	15.75	0.41	0.796	0.040	92	1.274	0.040	15.93	0.74	0.778	0.058
44	1.214	0.029	15.07	0.59	0.740	0.050	93	-	-	-	-	-	-
45	1.186	0.028	16.30	0.52	0.808	0.049	94	1.265	0.022	16.22	0.38	0.703	0.039
46	1.197	0.023	14.62	0.41	0.777	0.037	95	1.215	0.017	15.16	0.32	0.778	0.026
47	1.126	0.020	14.59	0.45	0.764	0.042	96	-	-	-	-	-	-
48	1.257	0.020	16.59	0.41	0.827	0.047	97	1.244	0.024	16.41	0.47	0.805	0.047
49	-	-	-	-	-	-	98	-	-	-	-	-	-

[a]Residue number. [b]Value given in (s^{-1}). [c]Error estimated from the repeated time points. [d]Error estimated from noise.

Table A.III. ^{15}N relaxation parameters of R40E-R98H *Pa*DDAH recorded at 500 MHz

# [a]	R_1 [b]	dR_1 [c]	R_2 [b]	dR_2 [c]	NOE	dNOE [d]	# [a]	R_1 [b]	dR_1 [c]	R_2 [b]	dR_2 [c]	NOE	dNOE [d]
99	1.229	0.029	15.14	0.58	0.818	0.046	148	1.214	0.015	15.52	0.31	0.773	0.028
100	1.155	0.017	15.28	0.33	0.775	0.034	149	1.243	0.025	15.09	0.47	0.790	0.044
101	-	-	-	-	-	-	150	-	-	-	-	-	-
102	1.301	0.049	14.95	0.81	0.730	0.039	151	1.195	0.033	14.71	0.58	0.718	0.047
103	1.280	0.023	14.67	0.39	0.757	0.039	152	1.231	0.030	15.23	0.60	0.810	0.046
104	1.152	0.016	14.95	0.30	0.769	0.033	153	-	-	-	-	-	-
105	1.161	0.034	14.59	0.66	0.812	0.060	154	1.150	0.023	14.95	0.47	0.822	0.047
106	1.214	0.024	15.23	0.43	0.772	0.039	155	1.240	0.034	14.60	0.52	0.773	0.053
107	-	-	-	-	-	-	156	1.174	0.019	13.68	0.33	0.651	0.038
108	1.174	0.029	15.25	0.62	0.813	0.055	157	-	-	-	-	-	-
109	1.262	0.029	14.28	0.52	0.825	0.044	158	1.333	0.036	13.82	0.61	0.524	0.050
110	-	-	-	-	-	-	159	1.214	0.045	14.74	0.81	0.619	0.054
111	1.263	0.044	15.63	0.80	0.807	0.056	160	1.080	0.016	12.62	0.29	0.415	0.028
112	1.272	0.037	15.64	0.72	0.825	0.058	161	1.232	0.028	15.31	0.54	0.683	0.046
113	-	-	-	-	-	-	162	1.241	0.035	17.69	0.68	0.720	0.050
114	1.199	0.050	15.25	0.94	0.740	0.057	163	1.166	0.064	16.34	1.36	0.731	0.092
115	1.187	0.057	15.72	0.99	0.741	0.068	164	1.281	0.051	16.28	1.03	0.822	0.063
116	1.312	0.058	17.06	1.30	0.762	0.064	165	1.282	0.039	17.91	0.75	0.791	0.052
117	1.274	0.035	15.93	0.71	0.749	0.044	166	1.242	0.042	16.32	0.88	0.821	0.064
118	1.258	0.032	15.91	0.65	0.920	0.054	167	1.219	0.026	15.99	0.58	0.788	0.045
119	1.202	0.048	15.05	0.89	0.789	0.063	168	1.199	0.018	15.89	0.38	0.766	0.035
120	1.171	0.055	14.10	0.98	0.801	0.075	169	1.237	0.054	15.06	1.01	0.779	0.068
121	1.198	0.040	16.01	0.81	0.788	0.057	170	1.235	0.052	15.24	0.94	0.727	0.060
122	1.273	0.038	15.85	0.79	0.775	0.059	171	-	-	-	-	-	-
123	1.171	0.031	14.48	0.66	0.766	0.052	172	1.222	0.050	14.80	0.92	0.801	0.065
124	-	-	-	-	-	-	173	-	-	-	-	-	-
125	-	-	-	-	-	-	174	-	-	-	-	-	-
126	1.218	0.030	15.20	0.65	0.794	0.050	175	-	-	-	-	-	-
127	1.187	0.038	15.66	0.71	0.792	0.056	176	1.252	0.042	15.35	0.79	0.703	0.052
128	1.341	0.064	14.81	1.21	0.851	0.084	177	-	-	-	-	-	-
129	1.193	0.028	14.58	0.54	0.718	0.046	178	1.177	0.025	15.66	0.48	0.756	0.047
130	1.208	0.030	15.02	0.58	0.722	0.050	179	1.265	0.024	12.87	0.38	0.697	0.040
131	1.271	0.040	14.36	0.61	0.750	0.049	180	1.418	0.079	14.04	1.22	0.684	0.068
132	1.193	0.034	14.71	0.73	0.782	0.063	181	1.321	0.026	16.50	0.44	0.661	0.037
133	1.174	0.035	16.10	0.67	0.778	0.054	182	-	-	-	-	-	-
134	1.235	0.033	15.33	0.62	0.811	0.051	183	1.216	0.024	14.00	0.43	0.764	0.042
135	1.251	0.026	15.51	0.55	0.775	0.047	184	-	-	-	-	-	-
136	1.237	0.027	16.03	0.50	0.810	0.049	185	-	-	-	-	-	-
137	1.302	0.042	16.55	0.85	0.842	0.061	186	1.219	0.027	15.40	0.52	0.780	0.046
138	1.273	0.016	16.09	0.32	0.803	0.027	187	1.193	0.022	15.46	0.42	0.758	0.034
139	1.220	0.023	16.03	0.44	0.812	0.041	188	1.227	0.020	15.35	0.39	0.796	0.041
140	-	-	-	-	-	-	189	1.239	0.021	16.63	0.41	0.788	0.036
141	1.240	0.025	15.21	0.47	0.793	0.042	190	1.211	0.023	15.42	0.41	0.795	0.043
142	1.229	0.023	14.57	0.41	0.658	0.018	191	1.232	0.028	15.30	0.57	0.826	0.050
143	-	-	-	-	-	-	192	1.124	0.020	14.24	0.43	0.796	0.044
144	-	-	-	-	-	-	193	1.221	0.038	15.98	0.77	0.802	0.066
145	1.269	0.034	18.03	0.72	0.838	0.057	194	-	-	-	-	-	-
146	1.238	0.028	16.86	0.56	0.736	0.041	195	1.192	0.036	15.83	0.72	0.726	0.063
147	1.213	0.020	16.24	0.42	0.760	0.040	196	-	-	-	-	-	-

[a]Residue number. [b]Value given in (s^{-1}). [c]Error estimated from the repeated time points. [d]Error estimated from noise.

Table A.III. ^{15}N relaxation parameters of R40E-R98H *Pa*DDAH recorded at 500 MHz

#[a]	R_1[b]	dR_1[c]	R_2[b]	dR_2[c]	NOE	dNOE[d]	#[a]	R_1[b]	dR_1[c]	R_2[b]	dR_2[c]	NOE	dNOE[d]
197	1.451	0.059	14.81	0.52	0.723	0.036	246	-	-	-	-	-	-
198	1.302	0.061	16.30	1.03	0.637	0.056	247	-	-	-	-	-	-
199	1.217	0.026	15.27	0.47	0.699	0.042	248	1.283	0.048	15.04	0.82	0.798	0.058
200	1.225	0.021	15.63	0.39	0.718	0.035	249	1.223	0.028	13.63	0.45	0.590	0.028
201	1.322	0.036	16.24	0.63	0.835	0.044	250	1.218	0.045	15.74	0.89	0.797	0.064
202	1.285	0.028	16.41	0.62	0.796	0.044	251	1.304	0.029	16.49	0.60	0.818	0.041
203	1.270	0.026	16.07	0.46	0.830	0.039	252	1.195	0.049	15.73	0.91	0.767	0.058
204	1.143	0.030	14.89	0.58	0.817	0.047	253	1.213	0.042	14.69	0.83	0.744	0.065
205	1.152	0.042	15.10	0.79	0.821	0.062	254	1.228	0.031	15.04	0.59	0.803	0.041
206	1.204	0.038	15.46	0.81	0.749	0.059							
207	1.212	0.047	15.89	0.99	0.762	0.062							
208	1.236	0.028	15.60	0.57	0.796	0.042							
209	-	-	-	-	-	-							
210	-	-	-	-	-	-							
211	1.230	0.027	16.42	0.57	0.796	0.045							
212	1.146	0.027	15.21	0.56	0.716	0.042							
213	-	-	-	-	-	-							
214	1.244	0.039	15.53	0.63	0.829	0.052							
215	-	-	-	-	-	-							
216	1.273	0.036	15.87	0.73	0.806	0.056							
217	1.230	0.045	14.00	0.82	0.797	0.065							
218	1.177	0.027	14.58	0.58	0.805	0.046							
219	-	-	-	-	-	-							
220	1.253	0.036	15.17	0.68	0.865	0.063							
221	1.189	0.047	15.72	0.94	0.803	0.071							
222	1.285	0.031	16.14	0.55	0.745	0.045							
223	1.259	0.024	16.85	0.49	0.766	0.041							
224	1.215	0.024	15.78	0.51	0.788	0.046							
225	1.202	0.019	16.33	0.38	0.730	0.033							
226	1.241	0.023	15.84	0.47	0.774	0.040							
227	1.202	0.016	16.03	0.30	0.759	0.034							
228	1.176	0.027	15.81	0.52	0.733	0.047							
229	1.203	0.029	15.28	0.57	0.773	0.046							
230	1.090	0.018	13.57	0.39	0.787	0.039							
231	1.076	0.020	13.97	0.41	0.765	0.038							
232	1.095	0.022	13.60	0.40	0.735	0.044							
233	1.177	0.033	15.31	0.65	0.801	0.054							
234	1.121	0.025	14.40	0.44	0.821	0.053							
235	1.213	0.028	14.84	0.53	0.787	0.047							
236	1.307	0.031	14.61	0.51	0.733	0.037							
237	1.254	0.039	14.61	0.66	0.723	0.054							
238	1.285	0.030	16.35	0.58	0.824	0.049							
239	1.268	0.042	17.13	0.74	0.768	0.060							
240	1.181	0.035	15.40	0.68	0.759	0.046							
241	1.310	0.050	14.85	0.82	0.800	0.060							
242	1.281	0.039	16.20	0.77	0.799	0.060							
243	1.302	0.041	15.44	0.72	0.792	0.055							
244	1.247	0.031	12.62	0.51	0.705	0.046							
245	1.250	0.035	15.41	0.62	0.742	0.045							

[a]Residue number. [b]Value given in (s^{-1}). [c]Error estimated from the repeated time points. [d]Error estimated from noise.

Table A.IV. ^{15}N relaxation parameters of R40E-R98H *Pa*DDAH recorded at 600 MHz

#[a]	R_1[b]	dR_1[c]	R_2[b]	dR_2[c]	NOE	dNOE[d]	#[a]	R_1[b]	dR_1[c]	R_2[b]	dR_2[c]	NOE	dNOE[d]
1	1.057	0.019	13.90	0.26	0.598	0.037	50	0.913	0.029	16.20	0.72	0.748	0.053
2	0.882	0.031	20.87	0.92	0.786	0.073	51	0.921	0.021	16.06	0.44	0.830	0.050
3	0.932	0.038	16.92	0.85	0.822	0.086	52	0.970	0.040	17.27	0.93	0.811	0.074
4	0.962	0.023	18.42	0.58	0.830	0.050	53	-	-	-	-	-	-
5	0.961	0.037	17.02	0.74	0.776	0.053	54	-	-	-	-	-	-
6	0.989	0.033	16.77	0.74	0.816	0.058	55	0.964	0.038	17.41	0.88	0.803	0.053
7	0.980	0.027	16.65	0.55	0.818	0.048	56	0.962	0.035	16.71	0.68	0.824	0.060
8	0.926	0.038	17.95	0.83	0.857	0.059	57	1.034	0.038	18.54	0.77	0.801	0.050
9	0.993	0.030	17.36	0.58	0.824	0.052	58	1.022	0.030	17.28	0.51	0.831	0.042
10	-	-	-	-	-	-	59	-	-	-	-	-	-
11	0.951	0.042	15.94	0.90	0.817	0.062	60	0.950	0.039	16.83	0.71	0.827	0.058
12	0.946	0.028	16.26	0.50	0.818	0.057	61	1.079	0.053	18.95	1.12	0.870	0.057
13	1.013	0.025	16.92	0.52	0.753	0.045	62	1.050	0.046	19.30	0.97	0.843	0.055
14	-	-	-	-	-	-	63	-	-	-	-	-	-
15	1.088	0.026	17.27	0.44	0.790	0.045	64	-	-	-	-	-	-
16	-	-	-	-	-	-	65	1.102	0.041	19.18	0.86	0.836	0.055
17	1.528	0.056	11.55	0.28	0.600	0.032	66	1.001	0.036	17.37	0.83	0.794	0.055
18	1.183	0.025	11.66	0.33	0.513	0.036	67	-	-	-	-	-	-
19	1.430	0.040	10.90	0.36	0.448	0.028	68	-	-	-	-	-	-
20	1.423	0.102	11.54	0.86	0.492	0.061	69	-	-	-	-	-	-
21	-	-	-	-	-	-	70	0.875	0.036	15.03	0.76	0.886	0.070
22	-	-	-	-	-	-	71	0.953	0.031	17.69	0.64	0.829	0.051
23	1.523	0.078	13.09	0.41	0.524	0.032	72	0.949	0.048	16.46	0.83	0.883	0.074
24	1.526	0.062	11.69	0.31	0.482	0.029	73	1.015	0.027	17.16	0.50	0.799	0.047
25	-	-	-	-	-	-	74	0.850	0.020	14.02	0.39	0.637	0.041
26	-	-	-	-	-	-	75	0.972	0.026	17.57	0.56	0.834	0.049
27	0.921	0.031	16.31	0.62	0.821	0.057	76	0.864	0.023	17.40	0.55	0.814	0.055
28	0.947	0.026	17.57	0.49	0.870	0.048	77	0.963	0.036	17.17	0.74	0.728	0.056
29	0.910	0.015	16.63	0.32	0.807	0.033	78	-	-	-	-	-	-
30	0.944	0.019	17.34	0.35	0.839	0.039	79	0.875	0.036	17.03	0.81	0.830	0.069
31	0.974	0.022	18.42	0.42	0.815	0.044	80	-	-	-	-	-	-
32	1.005	0.021	18.11	0.45	0.849	0.047	81	0.972	0.050	16.61	0.99	0.756	0.075
33	-	-	-	-	-	-	82	1.043	0.024	15.24	0.37	0.681	0.039
34	0.972	0.028	18.22	0.56	0.887	0.053	83	-	-	-	-	-	-
35	0.953	0.026	18.00	0.57	0.852	0.049	84	-	-	-	-	-	-
36	0.943	0.019	17.78	0.40	0.793	0.038	85	1.040	0.033	17.18	0.61	0.807	0.050
37	1.010	0.020	18.14	0.39	0.876	0.041	86	1.069	0.030	18.17	0.59	0.843	0.057
38	-	-	-	-	-	-	87	1.017	0.068	19.85	1.54	0.892	0.082
39	-	-	-	-	-	-	88	1.107	0.039	19.39	0.76	0.878	0.061
40	-	-	-	-	-	-	89	1.042	0.029	18.05	0.61	0.809	0.052
41	-	-	-	-	-	-	90	0.958	0.021	16.72	0.41	0.772	0.048
42	0.919	0.019	17.89	0.47	0.793	0.044	91	0.933	0.040	18.92	0.86	0.772	0.048
43	0.936	0.017	17.24	0.35	0.784	0.039	92	0.988	0.031	17.61	0.63	0.678	0.065
44	0.940	0.022	16.30	0.41	0.751	0.042	93	-	-	-	-	-	-
45	0.949	0.021	18.11	0.43	0.858	0.051	94	0.967	0.018	17.55	0.37	0.786	0.041
46	0.905	0.023	16.85	0.48	0.771	0.038	95	0.959	0.014	16.86	0.26	0.865	0.027
47	0.863	0.017	16.81	0.44	0.770	0.045	96	-	-	-	-	-	-
48	0.926	0.019	17.87	0.42	0.801	0.051	97	0.972	0.021	18.18	0.46	0.759	0.049
49	0.957	0.025	17.27	0.49	0.791	0.048	98	-	-	-	-	-	-

[a]Residue number. [b]Value given in (s^{-1}). [c]Error estimated from the repeated time points. [d]Error estimated from noise.

Table A.IV. ^{15}N relaxation parameters of R40E-R98H *Pa*DDAH recorded at 600 MHz

#[a]	R_1[b]	dR_1[c]	R_2[b]	dR_2[c]	NOE	dNOE[d]	#[a]	R_1[b]	dR_1[c]	R_2[b]	dR_2[c]	NOE	dNOE[d]
99	0.962	0.027	17.02	0.54	0.855	0.050	148	0.941	0.013	16.99	0.27	0.883	0.034
100	0.898	0.019	17.17	0.44	0.860	0.038	149	1.018	0.025	16.99	0.50	0.794	0.046
101	-	-	-	-	-	-	150	-	-	-	-	-	-
102	-	-	-	-	-	-	151	0.945	0.025	17.01	0.48	0.813	0.054
103	1.000	0.020	16.96	0.40	0.819	0.041	152	0.944	0.024	18.03	0.52	0.785	0.049
104	0.895	0.014	16.48	0.30	0.803	0.037	153	-	-	-	-	-	-
105	0.894	0.027	16.11	0.62	0.845	0.065	154	0.909	0.020	16.29	0.38	0.815	0.047
106	0.913	0.018	16.93	0.39	0.771	0.040	155	0.965	0.025	16.81	0.53	0.737	0.045
107	-	-	-	-	-	-	156	0.983	0.018	15.04	0.34	0.664	0.043
108	0.890	0.023	16.78	0.49	0.763	0.057	157	-	-	-	-	-	-
109	0.952	0.022	15.96	0.41	0.782	0.043	158	1.111	0.031	15.66	0.53	0.610	0.057
110	-	-	-	-	-	-	159	0.969	0.031	15.95	0.51	0.675	0.058
111	1.011	0.032	17.46	0.60	0.808	0.049	160	0.981	0.018	15.76	0.34	0.438	0.031
112	1.002	0.031	16.87	0.63	0.863	0.061	161	0.927	0.025	16.67	0.50	0.772	0.050
113	-	-	-	-	-	-	162	0.949	0.029	21.14	0.83	0.793	0.057
114	0.924	0.037	16.15	0.85	0.732	0.056	163	0.926	0.054	19.76	1.38	1.019	0.105
115	0.907	0.040	17.65	1.04	0.896	0.073	164	0.999	0.047	19.18	0.87	0.866	0.066
116	1.029	0.049	18.93	1.11	0.832	0.071	165	1.002	0.039	20.05	0.86	0.882	0.056
117	0.993	0.036	18.03	0.71	0.811	0.048	166	1.003	0.033	18.27	0.70	0.801	0.059
118	0.987	0.028	17.58	0.56	0.920	0.049	167	1.006	0.024	18.60	0.55	0.833	0.041
119	0.960	0.036	17.69	0.81	0.873	0.067	168	0.951	0.022	17.76	0.48	0.841	0.044
120	0.893	0.042	16.79	0.90	0.836	0.074	169	0.898	0.037	16.68	0.83	0.869	0.066
121	0.939	0.035	18.13	0.81	0.823	0.056	170	0.884	0.035	17.90	0.75	0.797	0.061
122	1.008	0.032	18.46	0.64	0.804	0.059	171	-	-	-	-	-	-
123	0.875	0.025	16.60	0.51	0.839	0.050	172	0.947	0.041	16.94	0.87	0.817	0.071
124	-	-	-	-	-	-	173	-	-	-	-	-	-
125	-	-	-	-	-	-	174	-	-	-	-	-	-
126	0.955	0.029	16.02	0.58	0.874	0.049	175	-	-	-	-	-	-
127	0.902	0.025	16.55	0.55	0.762	0.050	176	0.970	0.032	17.28	0.61	0.905	0.064
128	0.958	0.045	17.25	0.87	0.881	0.080	177						
129	0.899	0.024	16.15	0.51	0.773	0.051	178	0.927	0.020	16.83	0.39	0.818	0.048
130	1.014	0.030	17.31	0.61	0.790	0.058	179	1.030	0.022	14.38	0.31	0.662	0.040
131	0.968	0.031	16.62	0.56	0.770	0.054	180	1.079	0.057	16.42	0.96	0.786	0.081
132	0.871	0.029	16.30	0.76	0.820	0.081	181	1.030	0.018	17.00	0.36	0.795	0.044
133	0.960	0.038	16.34	0.76	0.770	0.059	182	-	-	-	-	-	-
134	1.014	0.031	17.38	0.62	0.859	0.051	183	0.923	0.017	16.29	0.34	0.793	0.042
135	0.935	0.019	17.21	0.39	0.839	0.046	184	-	-	-	-	-	-
136	0.937	0.018	17.37	0.41	0.783	0.041	185	-	-	-	-	-	-
137	0.948	0.032	17.21	0.73	0.874	0.061	186	0.925	0.020	16.12	0.43	0.817	0.049
138	1.034	0.016	18.26	0.34	0.850	0.033	187	0.925	0.018	16.53	0.39	0.849	0.037
139	0.962	0.018	17.88	0.42	0.829	0.038	188	1.006	0.021	17.28	0.41	0.807	0.036
140	-	-	-	-	-	-	189	0.953	0.015	16.48	0.35	0.802	0.041
141	0.922	0.023	17.55	0.51	0.914	0.049	190	0.970	0.020	17.19	0.43	0.820	0.044
142	0.929	0.018	16.77	0.41	0.881	0.044	191	0.925	0.026	17.36	0.56	0.806	0.050
143	-	-	-	-	-	-	192	0.854	0.019	15.67	0.40	0.807	0.044
144	0.967	0.023	17.79	0.50	0.835	0.045	193	0.914	0.031	17.04	0.64	0.860	0.068
145	0.959	0.027	20.13	0.68	0.827	0.055	194	-	-	-	-	-	-
146	0.999	0.022	19.57	0.51	0.843	0.046	195	0.942	0.028	17.12	0.56	0.770	0.056
147	0.941	0.016	18.03	0.33	0.782	0.041	196	-	-	-	-	-	-

[a]Residue number. [b]Value given in (s^{-1}). [c]Error estimated from the repeated time points. [d]Error estimated from noise.

Table A.IV. ^{15}N relaxation parameters of R40E-R98H *Pa*DDAH recorded at 600 MHz

#[a]	R_1[b]	dR_1[c]	R_2[b]	dR_2[c]	NOE	dNOE[d]	#[a]	R_1[b]	dR_1[c]	R_2[b]	dR_2[c]	NOE	dNOE[d]
197	1.198	0.058	16.42	0.43	0.752	0.046	246	-	-	-	-	-	-
198	1.050	0.040	17.82	0.76	0.738	0.062	247	-	-	-	-	-	-
199	0.989	0.026	17.37	0.54	0.741	0.044	248	0.961	0.035	18.12	0.82	0.896	0.064
200	0.957	0.018	16.66	0.36	0.836	0.038	249	1.006	0.028	13.98	0.41	0.740	0.036
201	1.042	0.025	17.98	0.46	0.931	0.053	250	0.932	0.035	18.08	0.72	0.821	0.054
202	0.978	0.022	18.15	0.46	0.854	0.046	251	1.048	0.025	17.90	0.52	0.845	0.042
203	1.054	0.026	17.81	0.44	0.833	0.040	252	0.938	0.037	17.29	0.69	0.787	0.059
204	0.952	0.026	17.51	0.57	0.797	0.043	253	0.891	0.033	15.93	0.63	0.837	0.072
205	0.888	0.027	17.37	0.69	0.824	0.057	254	0.978	0.024	17.62	0.56	0.853	0.044
206	0.972	0.039	17.14	0.69	0.818	0.051							
207	0.944	0.045	17.29	0.84	0.816	0.069							
208	0.943	0.024	17.41	0.53	0.853	0.042							
209	-	-	-	-	-	-							
210	-	-	-	-	-	-							
211	0.953	0.023	17.72	0.48	0.798	0.040							
212	0.927	0.020	16.48	0.47	0.771	0.043							
213	-	-	-	-	-	-							
214	0.975	0.029	16.49	0.62	0.825	0.048							
215	-	-	-	-	-	-							
216	0.977	0.028	17.95	0.65	0.870	0.057							
217	0.933	0.031	16.36	0.63	0.937	0.068							
218	0.931	0.022	16.21	0.49	0.877	0.049							
219	-	-	-	-	-	-							
220	0.909	0.025	17.27	0.58	0.786	0.054							
221	0.922	0.035	16.74	0.73	0.901	0.074							
222	0.968	0.020	17.91	0.52	0.792	0.046							
223	0.965	0.019	18.12	0.41	0.830	0.043							
224	1.002	0.022	17.93	0.42	0.783	0.045							
225	1.018	0.020	17.96	0.40	0.776	0.042							
226	0.956	0.018	17.96	0.40	0.834	0.043							
227	0.955	0.014	17.72	0.30	0.778	0.035							
228	0.903	0.017	17.47	0.38	0.787	0.046							
229	0.923	0.022	16.55	0.49	0.794	0.051							
230	0.932	0.017	15.55	0.35	0.813	0.043							
231	0.853	0.018	15.69	0.39	0.785	0.051							
232	0.852	0.015	15.44	0.35	0.768	0.043							
233	0.947	0.030	17.14	0.57	0.825	0.054							
234	0.861	0.018	14.89	0.40	0.857	0.050							
235	0.964	0.025	16.95	0.54	0.821	0.046							
236	1.037	0.026	16.72	0.47	0.867	0.044							
237	0.973	0.032	16.91	0.57	0.759	0.054							
238	0.994	0.026	18.08	0.51	0.846	0.048							
239	0.955	0.029	17.75	0.60	0.865	0.067							
240	0.906	0.031	17.10	0.68	0.798	0.047							
241	0.982	0.037	17.32	0.77	0.797	0.056							
242	1.102	0.035	17.71	0.65	0.821	0.062							
243	0.986	0.029	16.41	0.64	0.809	0.056							
244	1.022	0.025	16.03	0.45	0.724	0.040							
245	0.928	0.026	17.97	0.62	0.790	0.048							

[a]Residue number. [b]Value given in (s^{-1}). [c]Error estimated from the repeated time points. [d]Error estimated from noise.

Table A.V. ^{15}N relaxation parameters of R40E-R98H *Pa*DDAH recorded at 800 MHz

#[a]	R_1[b]	dR_1[c]	R_2[b]	dR_2[c]	NOE	dNOE[d]	#[a]	R_1[b]	dR_1[c]	R_2[b]	dR_2[c]	NOE	dNOE[d]
1	0.740	0.006	16.18	0.27	0.679	0.017	50	0.611	0.011	19.21	0.72	0.797	0.028
2	0.610	0.015	25.37	1.31	0.842	0.042	51	0.640	0.008	19.44	0.45	0.891	0.025
3	0.604	0.016	19.71	0.93	0.864	0.041	52	0.656	0.014	20.28	0.91	0.895	0.035
4	0.637	0.010	22.45	0.64	0.899	0.027	53	-	-	-	-	-	-
5	0.649	0.010	19.67	0.60	0.849	0.027	54	-	-	-	-	-	-
6	0.634	0.011	19.10	0.72	0.879	0.030	55	0.632	0.011	19.73	0.72	0.894	0.029
7	0.633	0.009	20.11	0.55	0.842	0.023	56	0.649	0.011	18.63	0.56	0.907	0.026
8	0.638	0.014	19.94	0.77	0.842	0.029	57	0.728	0.013	21.29	0.78	0.872	0.024
9	0.657	0.010	19.69	0.62	0.859	0.024	58	0.652	0.007	19.24	0.46	0.928	0.020
10	-	-	-	-	-	-	59	-	-	-	-	-	-
11	0.655	0.014	20.30	1.02	0.836	0.032	60	0.649	0.012	20.86	0.84	0.842	0.026
12	0.692	0.010	19.54	0.62	0.778	0.026	61	0.676	0.015	20.96	1.09	0.836	0.028
13	0.721	0.010	19.76	0.64	0.848	0.023	62	0.702	0.013	22.15	0.93	0.831	0.028
14	0.693	0.008	21.35	0.42	0.873	0.015	63	0.685	0.009	21.24	0.63	0.849	0.021
15	0.748	0.009	21.78	0.64	0.801	0.020	64	-	-	-	-	-	-
16	0.759	0.005	18.74	0.32	0.760	0.017	65	0.745	0.012	23.52	0.69	0.820	0.021
17	0.994	0.009	14.47	0.23	0.634	0.013	66	0.636	0.011	21.56	0.85	0.888	0.024
18	0.978	0.010	17.31	0.48	0.606	0.016	67	-	-	-	-	-	-
19	1.056	0.013	14.00	0.28	0.545	0.012	68	-	-	-	-	-	-
20	1.105	0.033	15.68	0.86	0.480	0.021	69	0.677	0.007	19.89	0.46	0.810	0.022
21	-	-	-	-	-	-	70	0.596	0.013	18.91	0.82	0.937	0.033
22	-	-	-	-	-	-	71	0.609	0.008	19.59	0.49	0.862	0.024
23	0.965	0.011	14.27	0.28	0.593	0.015	72	0.689	0.019	19.83	0.85	0.853	0.033
24	1.000	0.009	14.06	0.25	0.580	0.012	73	0.686	0.009	19.80	0.45	0.787	0.021
25	0.685	0.004	17.48	0.18	0.695	0.010	74	0.571	0.007	16.78	0.33	0.692	0.021
26	-	-	-	-	-	-	75	0.654	0.008	20.03	0.48	0.852	0.025
27	0.628	0.009	18.94	0.56	0.843	0.027	76	0.615	0.009	19.46	0.52	0.887	0.025
28	0.674	0.009	20.91	0.52	0.870	0.025	77	0.622	0.012	20.89	0.69	0.882	0.031
29	0.646	0.005	20.41	0.25	0.868	0.017	78	0.644	0.004	21.27	0.28	0.834	0.012
30	0.642	0.005	20.97	0.33	0.841	0.019	79	0.601	0.015	20.15	0.98	0.916	0.036
31	0.678	0.007	21.17	0.46	0.873	0.022	80	-	-	-	-	-	-
32	0.679	0.007	21.58	0.43	0.880	0.023	81	0.737	0.018	18.58	0.92	0.687	0.030
33	0.682	0.004	21.97	0.31	0.782	0.012	82	0.749	0.009	17.45	0.39	0.802	0.017
34	0.645	0.008	21.91	0.58	0.846	0.023	83	-	-	-	-	-	-
35	0.672	0.009	21.54	0.50	0.870	0.022	84	0.747	0.050	20.44	1.57	0.807	0.044
36	0.630	0.005	20.74	0.34	0.835	0.018	85	0.716	0.011	20.54	0.66	0.856	0.022
37	0.655	0.005	21.35	0.37	0.869	0.018	86	0.684	0.009	21.42	0.63	0.897	0.023
38	-	-	-	-	-	-	87	0.723	0.025	21.35	1.27	0.758	0.030
39	0.649	0.008	21.22	0.52	0.872	0.024	88	0.729	0.011	21.58	0.71	0.877	0.024
40	-	-	-	-	-	-	89	0.723	0.009	21.57	0.59	0.885	0.022
41	0.664	0.004	21.33	0.24	0.799	0.015	90	0.658	0.007	19.97	0.43	0.794	0.019
42	0.662	0.008	21.05	0.43	0.856	0.021	91	0.697	0.017	24.95	1.56	0.729	0.036
43	0.649	0.006	19.56	0.37	0.862	0.019	92	0.722	0.011	21.55	0.80	0.801	0.026
44	0.693	0.008	20.72	0.46	0.821	0.022	93	0.687	0.005	20.60	0.42	0.857	0.020
45	0.666	0.007	20.38	0.45	0.797	0.020	94	0.701	0.007	21.09	0.42	0.805	0.019
46	0.670	0.007	19.28	0.42	0.818	0.020	95	0.625	0.004	19.37	0.26	0.909	0.014
47	0.605	0.007	20.01	0.45	0.831	0.022	96	0.689	0.005	20.90	0.41	0.849	0.019
48	0.646	0.008	21.08	0.56	0.904	0.026	97	0.656	0.006	21.91	0.45	0.873	0.021
49	0.636	0.008	20.55	0.57	0.866	0.026	98	0.660	0.007	19.27	0.40	0.927	0.023

[a]Residue number. [b]Value given in (s^{-1}). [c]Error estimated from the repeated time points. [d]Error estimated from noise.

Table A.V. ^{15}N relaxation parameters of R40E-R98H *Pa*DDAH recorded at 800 MHz

#[a]	R_1[b]	dR_1[c]	R_2[b]	dR_2[c]	NOE	dNOE[d]	#[a]	R_1[b]	dR_1[c]	R_2[b]	dR_2[c]	NOE	dNOE[d]
99	0.632	0.008	20.45	0.49	0.864	0.024	148	0.628	0.006	19.69	0.35	0.885	0.022
100	0.611	0.006	19.47	0.38	0.858	0.020	149	0.668	0.008	20.79	0.43	0.847	0.022
101	-	-	-	-	-	-	150	0.659	0.005	20.71	0.32	0.768	0.014
102	0.666	0.015	18.58	0.55	0.807	0.029	151	0.670	0.010	20.49	0.51	0.839	0.022
103	0.647	0.005	19.25	0.33	0.882	0.019	152	0.633	0.006	20.39	0.39	0.857	0.022
104	0.627	0.004	19.53	0.24	0.823	0.015	153	0.678	0.008	20.07	0.44	0.898	0.019
105	0.607	0.009	19.26	0.57	0.877	0.031	154	0.615	0.006	19.65	0.33	0.872	0.020
106	0.639	0.006	19.95	0.32	0.824	0.019	155	0.635	0.009	19.43	0.51	0.875	0.024
107	0.623	0.010	19.69	0.43	0.819	0.029	156	0.696	0.006	17.94	0.32	0.696	0.019
108	0.650	0.009	20.02	0.45	0.826	0.024	157	-	-	-	-	-	-
109	0.681	0.008	19.70	0.34	0.844	0.020	158	0.842	0.016	17.56	0.53	0.645	0.026
110	-	-	-	-	-	-	159	0.701	0.012	18.57	0.56	0.730	0.027
111	0.727	0.013	21.14	0.64	0.893	0.025	160	0.747	0.007	18.66	0.31	0.535	0.017
112	0.679	0.010	20.54	0.58	0.870	0.025	161	0.643	0.009	19.57	0.53	0.817	0.025
113	0.672	0.004	20.16	0.23	0.849	0.016	162	0.677	0.011	25.33	0.87	0.791	0.027
114	0.618	0.012	19.31	0.76	0.886	0.029	163	0.634	0.024	20.88	1.48	0.842	0.045
115	0.664	0.018	19.87	1.06	0.893	0.034	164	0.693	0.014	23.01	1.01	0.761	0.031
116	0.735	0.020	23.75	1.50	0.861	0.028	165	0.690	0.013	24.00	0.84	0.891	0.027
117	0.693	0.010	21.90	0.73	0.899	0.023	166	0.692	0.013	21.43	0.80	0.851	0.029
118	0.682	0.008	21.13	0.50	0.902	0.021	167	0.711	0.011	21.88	0.65	0.809	0.025
119	0.685	0.014	20.54	0.75	0.889	0.030	168	0.644	0.007	19.97	0.46	0.851	0.021
120	0.648	0.017	20.12	1.07	0.887	0.034	169	0.624	0.015	19.01	0.76	0.884	0.031
121	0.642	0.011	21.00	0.72	0.823	0.026	170	0.679	0.011	20.23	0.70	0.895	0.027
122	0.666	0.010	21.04	0.66	0.856	0.027	171	-	-	-	-	-	-
123	0.624	0.008	19.81	0.52	0.902	0.024	172	0.658	0.013	19.57	0.85	0.893	0.032
124	0.697	0.008	22.21	0.46	0.816	0.017	173	0.651	0.010	20.27	0.62	0.850	0.025
125	0.705	0.008	21.35	0.56	0.863	0.016	174	0.662	0.005	21.07	0.32	0.904	0.014
126	0.636	0.011	19.06	0.54	0.830	0.025	175	0.619	0.010	20.25	0.56	0.887	0.025
127	0.636	0.010	19.08	0.62	0.837	0.025	176	0.654	0.011	20.49	0.57	0.867	0.026
128	0.664	0.016	20.20	0.90	0.903	0.036	177	-	-	-	-	-	-
129	0.677	0.009	18.23	0.47	0.774	0.022	178	0.654	0.007	19.68	0.41	0.892	0.022
130	0.645	0.010	20.22	0.67	0.805	0.022	179	0.792	0.009	16.93	0.32	0.730	0.018
131	0.695	0.012	19.12	0.61	0.812	0.026	180	0.733	0.023	19.00	0.80	0.797	0.033
132	0.666	0.015	19.78	1.01	0.831	0.033	181	0.767	0.010	21.41	0.49	0.745	0.020
133	0.623	0.011	18.89	0.67	0.846	0.027	182	-	-	-	-	-	-
134	0.633	0.008	20.03	0.54	0.866	0.024	183	0.668	0.006	19.92	0.33	0.795	0.017
135	0.655	0.007	20.39	0.39	0.858	0.020	184	0.690	0.005	19.95	0.30	0.820	0.017
136	0.656	0.006	21.53	0.33	0.836	0.019	185	-	-	-	-	-	-
137	0.665	0.010	22.08	0.71	0.916	0.029	186	0.680	0.009	20.02	0.47	0.839	0.023
138	0.668	0.005	21.39	0.33	0.821	0.015	187	0.674	0.007	21.11	0.35	0.840	0.019
139	0.643	0.006	21.17	0.37	0.847	0.019	188	0.670	0.006	19.25	0.32	0.861	0.020
140	-	-	-	-	-	-	189	0.680	0.006	20.04	0.32	0.854	0.018
141	0.647	0.008	20.96	0.53	0.911	0.022	190	0.684	0.006	20.36	0.30	0.854	0.018
142	0.663	0.007	21.42	0.50	0.830	0.022	191	0.657	0.010	19.97	0.47	0.873	0.025
143	0.670	0.004	21.71	0.25	0.828	0.017	192	0.598	0.007	18.19	0.32	0.835	0.022
144	0.671	0.007	20.99	0.43	0.902	0.024	193	0.638	0.010	20.25	0.67	0.879	0.032
145	0.667	0.010	25.08	0.71	0.898	0.028	194	0.607	0.005	19.19	0.28	0.844	0.019
146	0.647	0.008	23.25	0.57	0.807	0.022	195	0.680	0.009	20.93	0.63	0.794	0.026
147	0.682	0.006	21.15	0.29	0.836	0.017	196	-	-	-	-	-	-

[a]Residue number. [b]Value given in (s^{-1}). [c]Error estimated from the repeated time points. [d]Error estimated from noise.

Table A.V. ^{15}N relaxation parameters of R40E-R98H *Pa*DDAH recorded at 800 MHz

#[a]	R_1[b]	dR_1[c]	R_2[b]	dR_2[c]	NOE	dNOE[d]	#[a]	R_1[b]	dR_1[c]	R_2[b]	dR_2[c]	NOE	dNOE[d]
197	0.669	0.008	18.59	0.32	0.799	0.020	246	-	-	-	-	-	-
198	0.738	0.018	20.50	0.59	0.748	0.026	247	0.651	0.014	19.03	0.76	0.846	0.030
199	0.670	0.008	20.35	0.41	0.835	0.022	248	0.673	0.014	21.79	0.92	0.893	0.029
200	0.702	0.008	19.57	0.39	0.834	0.019	249	0.684	0.009	21.47	0.61	0.894	0.021
201	0.742	0.009	21.69	0.51	0.890	0.020	250	0.662	0.015	21.24	0.92	0.844	0.027
202	0.710	0.008	21.39	0.48	0.852	0.019	251	0.715	0.007	23.04	0.53	0.882	0.017
203	0.669	0.007	20.70	0.34	0.899	0.018	252	0.630	0.011	19.42	0.75	0.826	0.029
204	0.604	0.009	20.15	0.49	0.884	0.025	253	0.607	0.012	17.90	0.64	0.912	0.030
205	0.590	0.009	19.43	0.57	0.854	0.024	254	0.706	0.009	20.97	0.50	0.818	0.018
206	0.647	0.011	20.58	0.70	0.821	0.027							
207	0.615	0.013	19.57	0.87	0.862	0.030							
208	0.649	0.009	20.49	0.51	0.838	0.022							
209	-	-	-	-	-	-							
210	-	-	-	-	-	-							
211	0.654	0.008	21.87	0.52	0.844	0.021							
212	0.620	0.007	19.10	0.39	0.827	0.021							
213	0.633	0.009	20.06	0.62	0.909	0.028							
214	0.719	0.011	21.08	0.61	0.863	0.025							
215	-	-	-	-	-	-							
216	0.652	0.009	20.85	0.60	0.861	0.023							
217	0.668	0.012	19.00	0.47	0.867	0.026							
218	0.620	0.007	18.64	0.38	0.844	0.020							
219	-	-	-	-	-	-							
220	0.643	0.012	21.04	0.78	0.911	0.030							
221	0.625	0.013	20.71	0.80	0.883	0.035							
222	0.688	0.008	22.19	0.48	0.866	0.021							
223	0.658	0.006	20.85	0.37	0.857	0.019							
224	0.673	0.006	21.11	0.36	0.883	0.020							
225	0.652	0.007	20.29	0.43	0.880	0.024							
226	0.664	0.006	21.01	0.40	0.828	0.020							
227	0.709	0.005	20.94	0.29	0.850	0.017							
228	0.644	0.007	20.52	0.38	0.795	0.022							
229	0.682	0.009	19.65	0.46	0.878	0.023							
230	0.641	0.006	18.58	0.34	0.752	0.021							
231	0.612	0.006	17.70	0.31	0.814	0.020							
232	0.597	0.005	18.20	0.25	0.820	0.020							
233	0.621	0.009	20.06	0.53	0.858	0.024							
234	0.585	0.005	18.20	0.31	0.860	0.021							
235	0.645	0.008	19.97	0.47	0.867	0.024							
236	0.742	0.009	19.99	0.43	0.826	0.019							
237	0.656	0.010	18.47	0.56	0.767	0.023							
238	0.713	0.009	21.74	0.54	0.872	0.020							
239	0.713	0.013	21.53	0.83	0.838	0.028							
240	0.646	0.010	20.12	0.70	0.842	0.023							
241	0.696	0.013	21.29	0.75	0.878	0.026							
242	0.735	0.011	21.35	0.68	0.866	0.025							
243	0.655	0.011	20.49	0.66	0.895	0.028							
244	0.697	0.008	18.79	0.42	0.802	0.020							
245	0.684	0.011	21.23	0.63	0.837	0.021							

[a]Residue number. [b]Value given in (s^{-1}). [c]Error estimated from the repeated time points. [d]Error estimated from noise.

References

Achan V., Tran C.T., Arrigoni F., Whitley G.S., Leiper J.M. and Vallance P. (2002). All-*trans*-retinoic acid increases nitric oxide synthesis by endothelial cells: a role for the induction of dimethylarginine dimethylaminohydrolase. *Circ. Res.*, **90**, 764-769.

Achan V., Broadhead M., Malaki M., Whitley G., Leiper J. MacAllister R. and Vallance P. (2003). Asymmetric dimethylarginine causes hypertension and cardiac dysfunction in humans and is actively metabolized by dimethylarginine dimethylaminohydrolase. *Arterioscler. Thromb. Vasc. Biol.*, **23**, 1455-1459.

Adak S., Aulak K.S. and Stuehr D.J. (2002a). Direct evidence for nitric oxide production by a nitric-oxide synthase-like protein from *Bacillus subtilis*. *J. Biol. Chem.*, **277**, 16167-16171.

Adak S., Bilwes A.M., Panda K., Hosfield D., Aulak K.S., McDonald J.F., Tainer J.A., Getzoff E.D., Crane B.R. and Stuehr D.J. (2002b). Cloning, expression, and characterization of a nitric oxide synthase protein from *Deinococcus radiodurans*. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 107-112.

Akke M. (2002). NMR methods for characterizing microsecond to millisecond dynamics in recognition and catalysis. *Curr. Opin. Struct. Biol.*, **12**, 642-647.

Alderton W.K., Cooper C.E. and Knowles R.G. (2001). Nitric oxide synthases: structure, function and inhibition. *Biochem. J.*, **357**, 593-615.

Anglister J., Grzesiek S., Ren H., Klee C.B. and Bax A. (1993). Isotope-edited multidimensional NMR of calcineurin B in the presence of the non-deuterated detergent CHAPS. *J. Biomol. NMR*, **3**, 121-126.

Arita K., Hashimoto H., Shimizu T., Nakashima K., Yamada M. and Sato M. (2004). Structural basis for Ca^{2+}-induced activation of human PAD4. *Nat. Struct. Mol. Biol.*, **11**, 777-783.

Azuma H., Sato J., Hamasaki H., Sugimoto A., Isotani E. and Obayashi M.. (1995). Accumulation of endogenous inhibitors of nitric oxide synthesis and decreased content of L-arginine in regenerated endothelial cells. *Br. J. Pharmacol.*, **115**, 1001-1004.

Bahadur R.P., Chakrabarti P., Rodier F. and Janin J. (2004). A dissection of specific and non-specific protein-protein interfaces. *J. Mol. Biol.*, **336**, 943-955.

Bannister A.J. and Kouzarides T. (2005). Reversing histone methylation. *Nature*, **436**, 1103-1106.

Bannister A.J., Schneider R. and Kouzarides T. (2002). Histone methylation: dynamic or static? *Cell*, **109**, 801-806.
Beach H., Cole R., Gill M.L. and Loria J.P. (2005). Conservation of μs-ms enzyme motions in the apo- and substrate-mimicked state. *J. Am. Chem. Soc.*, **127**, 9167-9176.

Bedford M.T. and Richard S. (2005). Arginine methylation: an emerging regulator of protein function. *Mol. Cel.*, **18**, 263-272.

Bennett M.J., Choe S. and Eisenberg D. (1994) Domain swapping: entangling alliances between proteins. *Proc. Natl. Acad. Sci. U.S.A.*, **91**, 3127-3131.

Bennett M.J., Schlunegger M.P. and Eisenberg D. (1995). 3D domain swapping: a mechanism for oligomer assembly. *Protein Sci.*, **4**, 2455-2468.

Bogan A.A. and Thorn K.S. (1998). Anatomy of hot spots in protein interfaces. *J. Mol. Biol.*, **280**, 1-9.

Böger R.H., Bode-Böger S.M. and Frolich J.C. (1996). The L-arginine-nitric oxide pathway: role in atherosclearosis and therapeutic implications. *Artherosclerosis*, **127**, 1-11.

Böger R.H., Bode-Böger S.M., Thiele W., Junker W., Alexander K. and Frolich J.C. (1997). Biochemical evidence for impaired nitric oxide synthesis in patients with peripheral arterial occlusive disease. *Circulation*, **95**, 2068-2074.

Böger R.H., Bode-Böger S.M., Szuba A., Tsao P.S., Chan J.R., Tangphao O., Blaschke T.F. and Cooke J.P. (1998). Asymmetric dimethylarginine (ADMA): a novel risk factor for endothelial dysfunction: its role in hypercholesterolemia. *Circulation*, **98**, 1842-1847.

Böger R.H., Sydow K., Borlak J., Thum T., Lenzen H., Schubert B., Tsikas D. and Bode-Böger S.M. (2000). LDL cholesterol upregulates synthesis of asymmetrical dimethylarginine in human endothelial cells: involvement of S-adenosylmethionine-dependent methyltransferases. *Circ. Res.*, **87**, 99-105.

Böger R.H. and Bode-Böger S.M. (2001). The clinical pharmacology of L-arginine. *Annu. Rev. Pharmacol.*, **41**, 79-99.

Böger R.H. (2004). Asymmetric dimethylarginine, an endogenous inhibitor of nitric oxide synthase, explains the "L-arginine paradox" and acts as a novel cardiovascular risk factor. *J. Nutr.*, **134**, 2842S-2847S.

Böger R.H. and Ron E.S. (2005). L-Arginine improves vascular function by overcoming deleterious effects of ADMA, a novel cardiovascular risk factor. *Altern. Med. Rev.*, **10**, 14-23.

Bogle R.G., MacAllister R.J., Whitley G.S. and Vallance P. (1995). Induction of N^{η}-monomethyl-L-arginine uptake: a mechanism for differential inhibition of NO synthases? *Am. J. Physiol.*, **269**, 750-756.

Borchert T.V., Abagyan R., Jaenicke R. and Wierenga R.K. (1994). Design, creation, and characterization of a stable, monomeric triosephosphate isomerase. *Proc. Natl. Acad. Sci. U.S.A.*, **91**, 1515-1518.

Bracken C., Carr P.A., Cavanagh J. and Palmer A.G. III. (1999). Temperature dependence of intramolecular dynamics of the basic leucine zipper of GCN4: implications for the entropy of association with DNA. *J. Mol. Biol.*, **285**, 2133-2146.

Brahms H., Raymackers J., Union A., de Keyser F., Meheus L. and Luhrmann R. (2000). The C-terminal RG dipeptide repeats of the spliceosomal Sm proteins D1 and D3 contain symmetrical dimethylarginines, which form a major B-cell epitope for anti-Sm autoantibodies. *J.Biol. Chem.*, **275**, 16030-16036.

Brünger A.T. (1992). X-PLOR Version 3.1: a system for X-ray crystallography and NMR. Yale University Press, New Haven, CT, U.S.A.

Butterwick J.A., Loria J.P., Astrof N.S., Kroenke C.D., Cole R., Rance M. and Palmer A.G. III. (2004). Multiple time scale backbone dynamics of homologous thermophilic and mesophilic ribonuclease HI enzymes. *J. Mol. Biol.*, **339**, 855-871.

Cai H. and Harrison D.G. (2000). Endothelial dysfunction in cardiovascular diseases: the role of oxidant stress. *Circ. Res.*, **87**, 840-844.

Cavanagh J., Fairbrother W.J., Palmer A.G. III and Skelton N.J. (1995). Protein NMR Spectroscopy: Principles and Practice. Academic Press Inc., San Diego, U.S.A.

Cheng X., Collins R.E. and Zhang X. (2005). Structural and sequence motifs of protein (histone) methylation enzymes. *Annu. Rev. Biophys. Struct.*, **34**, 267-294.

Chothia C. and Janin J. (1975). Principles of protein-protein recognition. *Nature*, **256**, 705-708.

Clackson T. and Wells J.A. (1995). A hot spot of binding energy in a hormone-receptor interface. *Science*, **267**, 383-386.

Clarke S. (1993). Protein methylation. *Curr. Opin. Cell Biol.*, **5**, 977-983.

Clore G.M., Szabo A., Bax A., Kay L.E., Driscoll P.C. and Gronenborn A.M. (1990). Deviations from the simple two-parameter model-free approach to the interpretation of ^{15}N nuclear magnetic-relaxation of proteins. *J. Am. Chem. Soc.*, **112**, 4989-4991.

Clore G.M. and Gronenborn A.M. (1998). Determining the structures of large proteins and protein complexes by NMR. *Trends Biotechnol.*, **16**, 22-34.

Clore G.M., Gronenborn A.M. and Bax A. (1998). A robust method for determining the magnitude of the fully asymmetric alignment tensor of oriented macromolecules in the absence of structural information. *J. Magn. Reson.*, **133**, 216-221.

Cooke J.P., Andon N.A., Girerd X.J., Hirsch A.T. and Creager M.A. (1991). Arginine restores cholinergic relaxation of hypercholesterolemic rabbit thoracic aorta. *Circulation*, **83**, 1057-1062.

Cox J.M., Davis C.A., Chan C., Jourden M.J., Jorjorian A.D., Brym M.J., Snider M.J., Borders C.L. and Edmiston P.L. (2003). Generation of an active monomer of rabbit muscle creatine kinase by site-directed mutagenesis: the effect of quaternary structure on catalysis and stability. *Biochemistry*, **42**, 1863-1871.

Cuthbert G.L., Daujat S., Snowden A.W., Erdjument-Bromage H., Hagiwara T., Yamada M., Schneider R., Gregory P.D., Tempst P., Bannister A.J. and Kouzarides T. (2004). Histone deimination antagonizes arginine methylation. *Cell*, **118**, 545-553.

D'Alessio G. (1999). The evolutionary transition from monomeric to oligomeric proteins: tools, the environment, hypotheses. *Prog. Biophys. Mol. Biol.*, **72**, 271-298.

Dayoub H., Achan V., Adimoolam S., Jacobi J., Stuehlinger M.C., Wang B.Y., Tsao P.S., Kimoto M., Vallance P., Patterson A.J. and Cooke J.P. (2003). Dimethylarginine dimethylaminohydrolase regulates nitric oxide synthesis: genetic and physiological evidence. *Circulation*, **108**, 3042-3047.

Del Cardayre S.B., Ribo M., Yokel E.M., Quirk D.J., Rutter W.J. and Raines R.T. (1995). Engineering ribonuclease A: production, purification and characterization of wild-type enzyme and mutants at Gln11. *Protein Eng.*, **8**, 261-273.

Delaglio F., Grzesiek S., Vuister G.W., Zhu G., Pfeifer J. and Bax A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277-293.

DeLano W.L. (2002). Unraveling hot spots in binding interfaces: progress and challenges. *Curr. Opin. Struct. Biol.*, **12**, 14-20.

Denman R.B. (2005). PAD: the smoking gun behind arginine methylation signaling? *Bioessays*, **27**, 242-246.

Dowling R.B., Newton R., Robichaud A., Cole P.J., Barnes P.J. and Wilson R. (1998). Effect of inhibition of nitric oxide synthase on *Pseudomonas aeruginosa* infection of respiratory mucosa in vitro. *Am. J. Respir. Cell. Mol. Biol.*, **19**, 950-958.

Eisenmesser E.Z., Millet O., Labeikovsky W., Korzhnev D.M., Wolf-Watz M., Bosco D.A., Skalicky J.J., Kay L.E. and Kern D. (2005). Intrinsic dynamics of an enzyme underlies catalysis. *Nature*, **438**, 117-121.

Fairbrother W.J., Liu J., Pisacane P.I., Sliwkowski M.X. and Palmer A.G. III. (1998). Backbone dynamics of the EGF-like domain of heregulin-α. *J. Mol. Biol.*, **279**, 1149-1161.

Farrow N.A., Zhang O., Szabo A., Torchia D.A. and Kay L.E. (1995). Spectral density function mapping using ^{15}N relaxation data exclusively. *J. Biomol. NMR*, **6**, 153-162.

Feeney J., Batchelor J.G., Albrand J.P. and Roberts G.C.K. (1979). Effects of intermediate exhange processes on the estimation of equilibrium-constants by NMR. *J. Magn. Reson.*, **33**, 519-529.

Fernandez C., Hilty C., Wider G., Guntert P. and Wüthrich K. (2004). NMR structure of the integral membrane protein OmpX. *J. Mol. Biol.*, **336**, 1211-1221.

Fernandez C. and Wider G. (2003). TROSY in NMR studies of the structure and function of large biological macromolecules. *Curr. Opin. Struct. Biol.*, **13**, 570-580.

Fersht A. (1999). Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding. Freeman W.H. & Co Ltd, NY, U.S.A.

Freedberg D.I., Ishima R., Jacob J., Wang Y.X., Kustanovich I., Louis J.M. and Torchia D. (2002). Rapid structural fluctuations of the free HIV protease flaps in solution: relationship to crystal structures and comparison with predictions of dynamics calculations. *Protein Sci.*, **11**, 221-232.

Fritsche E., Bergner A., Humm A., Piepersberg W. and Huber R. (1998). Crystal structure of L-arginine:inosamine-phosphate amidinotransferase StrB1 from *Streptomyces griseus*: an enzyme involved in streptomycin biosynthesis. *Biochemistry*, **37**, 17664-17672.

Fritsche E., Humm A. and Huber R. (1999). The ligand-induced structural changes of human L-Arginine:Glycine amidinotransferase. A mutational and crystallographic study. *J. Biol. Chem.*, **274**, 3026-3032.

Furnham N., Blundell T.L., DePristo M.A. and Terwilliger T.C. (2006). Is one solution good enough? *Nat. Struct. Mol. Biol.*, **13**, 184-185.

Galkin A., Kulakova L., Sarikaya E., Lim K., Howard A. and Herzberg O. (2004). Structural insight into arginine degradation by arginine deiminase, an antibacterial and parasite drug target. *J. Biol. Chem.*, **279**, 14001-14008.

Galkin A., Lu X., Dunaway-Mariano D. and Herzberg O. (2005). Crystal structures representing the Michaelis complex and the thiouronium reaction intermediate of *Pseudomonas aeruginosa* arginine deiminase. *J. Biol. Chem.*, **280**, 34080-34087.

Gardner K.H., Rosen M.K. and Kay L.E. (1997). Global folds of highly deuterated, methyl-protonated proteins by multidimensional NMR. *Biochemistry*, **36**, 1389-1401.

Gardner K.H. and Kay L.E. (1998). The use of ^{2}H, ^{13}C, ^{15}N multidimensional NMR to study the structure and dynamics of proteins. *Annu. Rev. Biophys. Biomol. Struct.*, **27**, 357-406.

Gardner K.H., Zhang X., Gehring K. and Kay L.E. (1998). Solution NMR studies of a 42 kDa *Escherichia coli* maltose binding protein β-cyclodextrin complex: Chemical shift assignments and analysis. *J. Am. Chem. Soc.*, **120**, 11738–11748.

Garrett D.S., Seok Y.J., Liao D.I., Peterkofsky A., Gronenborn A.M. and Clore G.M. (1997). Solution structure of the 30 kDa N-terminal domain of the enzyme I of the *Escherichia coli* phosphoenolpyruvate:sugar phosphotransferase system by multidimensional NMR. *Biochemistry*, **36**, 2517-2530.

Gary J.D. and Clarke S. (1998). RNA and protein interactions modulated by protein arginine methylation. *Prog. Nucleic Acid Res. Mol. Biol.*, **61**, 65-131.

Goodsell D.S. and Olson A.J. (2000). Structural symmetry and protein function. *Annu. Rev. Biophys. Biomol. Struct.*, **29**, 105-153.

Goto N.K., Gardner K.H., Mueller G.A., Willis R.C. and Kay L.E. (1999). A robust and cost-effective method for the production of Val, Leu, Ile (δ1) methyl-protonated ^{15}N-, ^{13}C-, ^{2}H-labeled proteins. *J. Biomol. NMR*, **13**, 369-374.

Green S.M., Gittis A.G., Meeker A.K. and Lattman E.E. (1995). One-step evolution of a dimer from a monomeric protein. *Nat. Struct. Biol.*, **2**, 746-751.

Grzesiek S. and Bax A. (1993). Amino acid type determination in the sequential assignment procedure of uniformly $^{13}C/^{15}N$-enriched proteins. *J. Biomol. NMR*, **3**, 185-204.

Grzesiek S., Anglister J., Ren H. and Bax A. (1993). ^{13}C line narrowing by ^{2}H decoupling in $^{2}H/^{13}C/^{15}N$-enriched proteins. Application to triple resonance 4D *J* connectivity of sequential amides. *J. Am. Chem. Soc.*, **115**, 4369-4370.

Günther U.L. and Schaffhausen B. (2002). NMRKIN: simulating line shapes from two-dimensional spectra of proteins upon ligand binding. *J. Biomol. NMR*, **22**, 201-209.

Günther U.L., Mittag T. and Schaffhausen B. (2002). Probing Src homology 2 domain ligand interactions by differential line broadening. *Biochemistry*, **41**, 11658-11669.

Hansen P.E. (1988). Isotope effects in nuclear shielding. *Prog. Nucl. Magn. Reson. Spec.*, **20**, 207-255.

Hendsch Z.S. and Tidor B. (1994). Do salt bridges stabilize proteins? A continuum electrostatic analysis. *Protein Sci.*, **3**, 211-226.

Hobbs A.J., Higgs A. and Moncada S. (1999). Inhibition of nitric oxide synthase as a potential therapeutic target. *Annu. Rev. Pharmacol. Toxicol.*, **39**, 191-220.

Holden D.P., Cartwright J.E., Nussey S.S. and Whitley G.S. (2003). Estrogen stimulates dimethylarginine dimethylaminohydrolase activity and the metabolism of asymmetric dimethylarginine. *Circulation*, **108**, 1575-1580.

Homans S.W. (2005). Probing the binding entropy of ligand-protein interactions by NMR. *Chembiochem*, **6**, 1585-1591.

Honig B. and Nicholls A. (1995). Classical electrostatics in biology and chemistry. *Science*, **268**, 1144-1149.

Hornak V., Okur A., Rizzo R.C. and Simmerling C. (2006). HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 915-920.

Hubbard S.J and Thornton J.M. (1993). NACCESS. Version 2.0. Computer program, Department of Biochemistry and Molecular Biology, University College London. http://www.biochem.ucl.ac.uk/~roman/naccess/naccess.html.

Humm A., Fritsche E., Steinbacher S. and Huber R. (1997). Crystal structure and mechanism of human L-arginine:glycine amidinotransferase: a mitochondrial enzyme involved in creatine biosynthesis. *EMBO J.*, **16**, 3373-85.

Huntley J.J., Scrofani S.D., Osborne M.J., Wright P.E. and Dyson H.J. (2000). Dynamics of the metallo-β-lactamase from *Bacteroides fragilis* in the presence and absence of a tight-binding inhibitor. *Biochemistry*, **39**, 13356-13364.

Hwang P.M., Choy W.Y., Lo E.I., Chen L., Forman-Kay J.D., Raetz C.R.H., Prive G.G., Bishop R.E. and Kay L.E. (2002). Solution structure and dynamics of the outer membrane enzyme PagP by NMR. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 13560-13565.

Ito A., Tsao P.S., Adimoolam S., Kimoto M., Ogawa T. and Cooke J.P. (1999). Novel mechanism for endothelial dysfunction: dysregulation of dimethylarginine dimethylaminohydrolase. *Circulation*, **99**, 3092-3095.

Jaenicke R.J. (2000). Stability and stabilization of globular proteins in solution. *Biotech.*, **79**, 193-203.

Jawad Z. and Paoli M. (2002). Novel sequences propel familiar folds. *Structure*, **10**, 447-454.

Johnson J.E and Speir J.A. (1997). Quasi-equivalent viruses: a paradigm for protein assemblies. *J. Mol. Biol.*, **269**, 665-675.

Jones L.C., Tran C.T., Leiper J.M., Hingorani A.D. and Vallance P. (2003). Common genetic variation in a basal promoter element alters DDAH2 expression in endothelial cells. *Biochem. Biophys. Res. Commun.*, **310**, 836-843.

Jones S. and Thornton J.M. (1996). Principles of protein-protein interactions. *Proc. Natl. Acad. Sci. U.S.A.*, **93**, 13-20.

Kakimoto Y. and Akazawa S. (1970). Isolation and identification of N^{η},N^{η}- and N^{η},$N^{\eta'}$-dimethylarginine, N^{ε}-mono-, di-, and trimethyllysine, and glucosylgalactosyl- and galactosyl-delta-hydroxylysine from human urine. *J. Biol. Chem.*, **245**, 5751-5758.

Katoh E., Louis J.M., Yamazaki T., Gronenborn A.M., Torchia D.A. and Ishima R. (2003). A solution NMR study of the binding kinetics and the internal dynamics of an HIV-1 protease-substrate complex. *Protein Sci.*, **12**, 1376-1385.

Katusic Z.S. (2001). Vascular endothelial dysfunction: does tetrahydrobiopterin play a role? *Am. J. Physiol. Heart Circ. Physiol.*, **281**, 981-986.

Kay L.E., Torchia D.A. and Bax, A. (1989). Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry*, **28**, 8972-8979.

Kay L.E., Keifer P. and Saarinen T. (1992a). Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc.*, **114**, 10663-10665.

Kay L.E., Nicholson L.K., Delaglio F., Bax A. and Torchia D.A. (1992b). Pulse sequences for removal of the effects of cross correlation between dipolar and chemical-shift anisotropy relaxation mechanisms on the measurement of heteronuclear T_1 and T_2 values in proteins. *J. Mag. Reson.*, **97**, 359-375.

Kay L.E. (2005). NMR studies of protein structure and dynamics. *J. Magn. Reson.*, **173**, 193-207.

Kearney P.L., Bhatia M., Jones N.G., Yuan L., Glascock M.C., Catchings K.L., Yamada M. and Thompson P.R. (2005). Kinetic characterization of protein arginine deiminase 4: a transcriptional corepressor implicated in the onset and progression of rheumatoid arthritis. *Biochemistry*, **44**, 10570-10582.

Keeler J. (2005). Understanding NMR spectroscopy. John Wiley & Sons Ltd., Hoboken, NJ, U.S.A.

Kempf J.G. and Loria J.P. (2004). Measurement of intermediate exchange phenomena. *Methods Mol. Biol.*, **278**, 185-231.

Kern D., Eisenmesser E.Z. and Wolf-Watz M. (2005). Enzyme dynamics during catalysis measured by NMR spectroscopy. *Methods Enzymol.*, **394**, 507-524.

Kielstein J.T. and Zoccali C. (2005). Asymmetric dimethylarginine: a cardiovascular risk factor and a uremic toxin coming of age? *Am. J. Kidney. Dis.*, **46**, 186-202.

Kimoto M., Miyatake S., Sasagawa T., Yamashita H., Okita M., Oka T., Ogawa T. and Tsuji H. (1998). Purification, cDNA cloning and expression of human N^{η},N^{η}-dimethylarginine dimethylaminohydrolase. *Eur. J. Biochem.*, **258**, 863-868.

Kimoto M., Tsuji H., Ogawa T. and Sasaoka K. (1993). Detection of N^{η},N^{η}-dimethylarginine dimethylaminohydrolase in the nitric oxide-generating systems of rats using monoclonal antibody. *Arch. Biochem. Biophys.*, **300**, 657-662.

Knipp M. and Vašák M. (2000). A colorimetric 96-well microtiter plate assay for the determination of enzymatically formed citrulline. *Anal. Biochem.*, **286**, 257-264.

Knipp M., Braun O. and Vašák M. (2005). Searching for DDAH inhibitors: S-nitroso-L-homocysteine is a chemical lead. *J. Am. Chem. Soc.*, **127**, 2372-2373.

Knowles R.G. and Moncada S. (1992) Nitric oxide as a signal in blood vessels. *Trends Biochem. Sci.*, **17**, 399-402.

Koradi R., Billeter M. and Wüthrich K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics.*, **14**, 51-55.

Kovrigin E.L. and Loria J.P. (2006). Enzyme dynamics along the reaction coordinate: critical role of a conserved residue. *Biochemistry*, **45**, 2636-2647.

Kraulis P.J. (1989). ANSIG - A program for the assignment of protein ^{1}H 2D-NMR spectra by interactive computer-graphics. *J. Magn. Reson.*, **84**, 627-633.

Kroenke C.D., Loria J.P., Lee L.K., Rance M. and Palmer A.G. III. (1998). Longitudinal and transverse ^{1}H-^{15}N dipolar/^{15}N chemical shift anisotropy relaxation interference: unambiguous determination of rotational diffusion tensors and chemical exchange effects in biological macromolecules. *J. Am. Chem. Soc.*, **120**, 7905-7915.

Kroenke C.D., Rance M. and Palmer A.G. III. (1999). Variability of the ^{15}N chemical shift anisotropy in *Escherichia coli* ribonuclease H in solution. *J. Am. Chem. Soc.*, **121**, 10119-10125.

Larsen T.A., Olson A.J. and Goodsell D.S. (1998). Morphology of protein-protein interfaces. *Structure*, **6**, 421-427.

Larsson G., Schleucher J., Onions J., Hermann S., Grundstrom T. and Wijmenga S.S. (2005). Backbone dynamics of a symmetric calmodulin dimer in complex with the calmodulin-binding

domain of the basic-helix-loop-helix transcription factor SEF2-1/E2-2: A highly dynamic complex. *Biophys. J.*, **89**, 1214-1226.

Laskowski R.A. (1995). SURFNET: a program for visualizing molecular surfaces, cavities, and intermolecular interactions. *J. Mol. Graph.*, **13**, 323-330.

Lee J., Sayegh J., Daniel J., Clarke S. and Bedford M.T. (2005). PRMT8, a new membrane-bound tissue-specific member of the protein arginine methyltransferase family. *J. Biol. Chem.*, **280**, 32890-32896.

Lee L.K., Rance M., Chazin W.J. and Palmer A.G. III. (1997). Rotational diffusion anisotropy of proteins from simultaneous analysis of ^{15}N and ^{13}C alpha nuclear spin relaxation. *J. Biomol. NMR*, **9**, 287-298.

Leiper J. and Vallance P. (1999). Biological significance of endogenous methylarginines that inhibit nitric oxide synthases. *Cardiovasc. Res.*, **43**, 542-548.

Leiper J.M., Parry H., Kimoto M., Ogawa T., Russell R.J., Whitley G.S.J. and Vallance P. (1999). Identification of two human dimethylarginine dimethylaminohydrolases with distinct tissue distributions and homology with microbial arginine deiminases. *Biochem. J.*, **343**, 209-214.

Leiper J., Murray-Rust J., McDonald N. and Vallance P. (2002). S-nitrosylation of dimethylarginine dimethylaminohydrolase regulates enzyme activity: further interactions between nitric oxide synthase and dimethylarginine dimethylaminohydrolase. *Proc. Natl. Acad. Sci. U. S. A.*, **99**, 13527-13532.

Leiper J.M., Nandi M., Malaki M., Murray-Rust J., O'Hara B. and Vallance P. (2006). Disruption of methylarginine metabolism impairs vascular homeostasis. *Submitted.*

Levitt M.H. (2001). Spin Dynamics. Basics of Nuclear Magnetic Resonance. John Wiley & Sons, Ltd, Hoboken, NJ, U.S.A.

Lian L.Y and Roberts G.C. (1993). Effects of chemical exchange on NMR spectra. *NMR of macromolecules: A practical approach,* Oxford University Press, U.K., 153-182.

Lipari G. and Szabo A. (1982a). Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules. 1. Theory and range of validity. *J. am. chem. soc.*, **17**, 4546-4559.

Lipari G. and Szabo A. (1982b). Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules. 2. Analysis of experimental results *J. Am. Chem. Soc.*, **104**, 4559-4570.

Loria J.P., Rance M. and Palmer A.G. III. (1999). A TROSY CPMG sequence for characterizing chemical exchange in large proteins. *J. Biomol. NMR*, **15**, 151-155.

Lu X., Li L., Wu R., Feng X., Li Z., Yang H.Y., Wang C.H., Guo H., Galkin A., Herzberg O., Mariano P.S., Martin B.M. and Dunaway-Mariano D. (2006). Kinetic analysis of *Pseudomonas aeruginosa* arginine deiminase mutants and alternate substrates provides insight into structural determinants of function. *Biochemistry*, **45**, 1162-1172.

Lukin J.A., Kontaxis G., Simplaceanu V., Yuan, Y., Bax A. and Ho C. (2003). Quaternary structure of hemoglobin in solution. *Proc. Natl. Acad. Sci. U. S. A.*, **100**, 517-520.

Luz Z. and Meiboom S. (1963). Nuclear magnetic resonance study of protolysis of trimethylammonium ion in aqueous solution – Order of the reaction with respect to solvent. *J. Chem. Phys.*, **39**, 366-370.

MacAllister R.J., Fickling S.A., Whitley G.S. and Vallance P. (1994). Metabolism of methylarginines by human vasculature; implications for the regulation of nitric oxide synthesis. *Br. J. Pharmacol.*, **112**, 43-48.

Magalhães B.S., Harris R., Plevin M.J. and Driscoll P.C. (2004). Backbone ^{1}H, ^{13}C, and ^{15}N resonance assignments for a 29 kD monomeric variant of *Pseudomonas aeruginosa* dimethylarginine dimethylaminohydrolase. *J. Biomol. NMR*, **29**, 463-464.

Majerski Z., Zuanic M. and Metelko B. (1985). Deuterium-isotope effects on ^{13}C chemical-shifts of protoadamantane. Evidence for geometrical dependence of $^{3}\Delta$ and $^{4}\Delta$ effects. *J. Am. Chem. Soc.*, **107**, 1721-1726.

Mandel A.M., Akke M. and Palmer, A.G. III. (1995). Backbone dynamics of *Escherichia coli* Ribonuclease HI: correlations with structure and function in an active enzyme. *J. Mol. Biol.* **246**, 144-163.

Marianayagam N.J., Sunde M. and Matthews J.M. (2004). The power of two: protein dimerization in biology. *Trends Biochem. Sci.*, **29**, 618-625.

Markus M.A., Dayie K.T., Matsudaira P. and Wagner G. (1994). Effect of deuteration on the amide proton relaxation rates in proteins. Heteronuclear NMR experiments on villin 14T. *J. Magn. Reson. B.*, **105**, 192-195.

McBride A.E. and Silver P.A. (2001). State of the Arg:Protein methylation at arginine comes of age. *Cell*, **106**, 5-8.

McDermott J.R. (1976). Studies on the catabolism of N^{η}-methylarginine, N^{η},N^{η}-dimethylarginine and $N^{\eta},N^{\eta'}$-dimethylarginine in the rabbit. *Biochem. J.*, **154**, 179-184.

McDonald I. K. and Thornton J.M. (1994). Satisfying hydrogen-bonding potential in proteins. *J. Mol. Biol.*, **238**, 777-793.

McElheny D., Schnell J.R., Lansing J.C., Dyson H.J. and Wright P.E. (2005). Defining the role of active-site loop fluctuations in dihydrofolate reductase catalysis. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 5032-5037.

McRorie D.K. and Voelker P.J. (1993). Self-associating Systems in the Analytical Ultracentrifuge. Beckman Instruments Inc., CA, U.S.A.

Millet O., Loria J.P., Kroenke C.D., Pons M. and Palmer A.G. III. (2000). The static magnetic field dependence of chemical exchange linebroadening defines the NMR chemical shift time scale. *J. Am. Chem. Soc.*, **122**, 2867-2877.

Mittag T., Schaffhausen B. and Günther U.L. (2003). Direct observation of protein-ligand interaction kinetics. *Biochemistry*, **42**, 11128-11136.

Mittag T., Schaffhausen B. and Günther U.L. (2004). Tracing kinetic intermediates during ligand binding. *J. Am. Chem. Soc.*, **126**, 9017-9023.

Miyazaki H., Matsuoka H., Cooke J.P., Usui M., Ueda S., Okuda S. and Imaizumi T. (1999). Endogenous nitric oxide synthase inhibitor: a novel marker of atherosclerosis. *Circulation*, **99**, 1141-1146.

Mulder F.A., Ayed A., Yang D., Arrowsmith C.H. and Kay L.E. (2000). Assignment of $^{1}H(N)$, ^{15}N, $^{13}C(\alpha)$, ^{13}CO and $^{13}C(\beta)$ resonances in a 67 kDa p53 dimer using 4D-TROSY NMR spectroscopy. *J. Biomol. NMR*, **18**, 173-176.

Mulder F.A., Skrynnikov N.R., Hon B., Dahlquist F.W. and Kay L.E. (2001). Measurement of slow (μs-ms) time scale dynamics in protein side chains by ^{15}N relaxation dispersion NMR spectroscopy: application to Asn and Gln residues in a cavity mutant of T4 lysozyme. *J. Am. Chem. Soc.*, **123**, 967-975.

Murray-Rust J., Leiper J., McAlister M., Phelan J., Tilley S., Maria J.S., Vallance P. and McDonald N. (2001). Structural insights into the hydrolysis of cellular nitric oxide synthase inhibitors by dimethylarginine dimethylaminohydrolase. *Nat. Struct. Biol.*, **8**, 679-683.

Nicholls A., Sharp K.A. and Honig B. (1991). Protein folding and association – insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins*, **11**, 281-296.

Nicolle L.E. (2002). Resistant pathogens in urinary tract infections. *J. Am. Geriatr. Soc.*, **50**, 230-235.

Ogawa T., Kimoto M. and Sasaoka K. (1989). Purification and properties of a new enzyme, N^{η},N^{η}-dimethylarginine dimethylaminohydrolase from rat kidney. *J. Biol. Chem.*, **264**, 10205-10209.

Ollerenshaw J.E., Tugarinov V., Skrynnikov N.R. and Kay L.E. (2005). Comparison of $^{13}CH_3$, $^{13}CH_2D$, and $^{13}CHD_2$ methyl labeling strategies in proteins. *J. Biomol. NMR*, **33**, 25-41.

Osanai T., Saitoh M., Sasaki S., Tomita H., Matsunaga T. and Okumura K. (2003). Effect of shear stress on asymmetric dimethylarginine release from vascular endothelial cells. *Hypertension*, **42**, 985-990.

Osborne M.J., Schnell J., Benkovic S.J., Dyson H.J. and Wright P.E. (2001). Backbone dynamics in dihydrofolate reductase complexes: role of loop flexibility in the catalytic mechanism. *Biochemistry*, **40**, 9846-9859.

Palmer A.G. III, Cavanagh J., Wright P.E. and Rance M. (1991a). Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy. *J. Magn. Reson.*, **93**, 151-170.

Palmer A. G. III, Rance M. and Wright P.E. (1991b). Intramolecular motions of a zinc finger DNA-binding domain from xfin characterized by proton-detected natural abundance ^{13}C heteronuclear NMR spectroscopy. *J. Am. Chem. Soc.*, **113**, 4371–4380.

Palmer A.G. III, Skelton N.J., Chazin W.J., Wright P.E. and Rance M. (1992). Suppression of the effects of cross-correlation between dipolar and anisotropic chemical shift relaxation mechanisms in the measurement of spin-spin relaxation rates. *Mol. Phys.*, **75**, 699-711.

Palmer A.G. III., Williams J. and McDermott A. (1996). Nuclear magnetic resonance studies of biopolymer dynamics. *J. Phys. Chem.*, **100**, 13293-13310.

Palmer A.G. III., Kroenke C.D. and Loria J.P. (2001). Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. *Methods Enzymol.*, **339**, 204-238.

Palmer A.G. III. (2004). NMR characterization of the dynamics of biomacromolecules. *Chem. Rev.*, **104**, 3623-3640.

Palmer R.M., Ashton D.S. and Moncada S. (1988). Vascular endothelial cells synthetize nitric oxide from L-arginine. *Nature*, **333**, 664-666.

Paoli M. (2001). An elusive propeller-like fold. *Nat. Struct. Biol.*, **8**, 744-745.

Pellecchia M. (2005). Solution nuclear magnetic resonance spectroscopy techniques for probing intermolecular interactions. *Chem. Biol.*, **12**, 961-971.

Pellecchia M., Sem D.S. and Wüthrich K. (2002). NMR in drug discovery. *Nat. Rev. Drug. Discov.*, **1**, 211-219.

Peng J.W. and Wagner G. (1992). Mapping of the spectral densities of N-H bond motions in eglin c using heteronuclear relaxation experiments. *Biochemistry*, **31**, 8571-8586.

Peng J.W. and Wagner G. (1994). Investigation of protein motions via relaxation measurements. *Methods Enzymol.*, **239**, 563-596.

Perkins S.J. (1986). Protein volumes and hydration effects. The calculations of partial specific volumes, neutron scattering matchpoints and 280-nm absorption coefficients for proteins and glycoproteins from amino acid sequences. *Eur. J. Biochem.*, **157**, 169-180.

Pervushin K., Riek R., Wider G. and Wüthrich K. (1997). Attenuated T_2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl. Acad. Sci. U.S.A.*, **94**, 12366-12371.

Pfuhl M., Chen H.A., Kristensen S.M. and Driscoll P.C. (1999). NMR exchange broadening arising from specific low affinity protein self-association: analysis of nitrogen-15 nuclear relaxation for rat CD2 domain1. *J. Biomol. NMR*, **14**, 307-320.

Phan I.Q., Boyd J. and Campbell I.D. (1996). Dynamic studies of a fibronectin type I module pair at three frequencies: anisotropic modelling and direct determination of conformational exchange. *J. Biomol. NMR*, **8**, 369-378.

Piotto M., Saudek V. and Sklenar V. (1992). Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR*, **2**, 661-665.

Plevin M.J. (2003). Multidimensional, heteronuclear NMR spectroscopy of high molecular weight proteins in solution: Application to a bacterial dimethylarginine dimethylaminohydrolase. Thesis submitted for the degree of Doctor of Philosophy, University of London.

Plevin M.J., Magalhães B.S., Harris R., Sankar A., Perkins S.J. and Driscoll P.C. (2004). Characterization and manipulation of the *Pseudomonas aeruginosa* dimethylarginine dimethylaminohydrolase monomer-dimer equilibrium. *J. Mol. Biol.*, **341**, 171-184.

Pollock J.S., Forstermann U., Mitchell J.A., Warner T.D., Schmidt H.H., Nakane M and Murad F. (1991). Purification and characterization of particulate endothelium-derived relaxing factor synthase from cultured and natuve bovine aortic endothelial cells. *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 10480-1484.

Ponstingl H., Kabir T., Gorse D. and Thornton J.M. (2005). Morphological aspects of oligomeric protein structures. *Prog. Biophys. Mol. Biol.*, **89**, 9-35.

Radkiewicz J.L. and Brooks C.L. (2000). Protein dynamics in enzymatic catalysis: Exploration of dihydrofolate reductase. *J. Am. Chem. Soc.*, **122**, 225-231.

Radomski M.W., Zahar T. and Salas E. (1996). Nitric oxide in platelets. *Methods Enzymol.*, **269**, 88-107.

Rajarathnam K., Sykes B.D., Kay C.M., Dewald B., Geiser T., Baggiolini M. and Clarklewis I. (1994). Neutrophil activation by monomeric interleukin-8. *Science*, **264**, 90-92.

Rees D.D., Palmer R.M.J. and Moncada S. (1989). Role of endothelium-derived nitric oxide in the regulation of blood pressure. *Proc. Natl. Acad. Sci. U.S.A.*, **86**, 3375-3378.

Reibarkh M., Malia T.J. and Wagner G. (2006). NMR distinction of single- and multiple-mode binding of small-molecule protein ligands. *J. Am. Chem. Soc.*, **128**, 2160-2161.

Riek R., Pervushin K. and Wüthrich K. (2000). TROSY and CRINEPT: NMR with large molecular and supramolecular structures in solution. *Trends Biochem. Sci.*, **25**, 462-468.

Rosen M.K., Gardner K.H., Willis R.C., Parris W.E., Pawson T. and Kay L.E. (1996). Selective methyl group protonation of perdeuterated proteins. *J. Mol. Biol.*, **263**, 627-636.

Rossiter S., Smith C.L., Malaki M., Nandi M., Gill H., Leiper J.M., Vallance P. and Selwood D.L. (2005). Selective substrate-based inhibitors of mammalian dimethylarginine dimethylaminohydrolase. *J. Med. Chem.*, **48**, 4670-4678.

Rule G.S. and Hitchens K. (2006). Fundamentals of Protein NMR Spectroscopy. Springer, Dordrecht, The Netherlands.

Sakurai K. and Goto Y. (2002). Manipulating monomer-dimer equilibrium of bovine β-lactoglobulin by amino acid substitution. *J. Biol. Chem.*, **277**, 25735-25740.

Salzmann M., Pervushin K., Wider G., Senn H. and Wüthrich K. (2000). NMR assignment and secondary structure determination of an octameric 110 kDa protein using TROSY in triple resonance experiments. *J. Am. Chem. Soc.*, **122,** 7543-7548.

Sambrook J. and Russell D. (2000). Molecular cloning (3-volume set). Cold Spring Harbor Laboratory Press, U.S.A.

Santa-Maria J., Vallance P., Charles I.G. and Leiper J.M. (1999). Identification of microbial dimethylarginine dimethylaminohydrolase enzymes. *Mol. Microbiol.*, **33**, 1278-1279.

Sattler M., Schleucher J. and Griesinger C. (1999). Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Spec.*, **34**, 93-158.

Schulga A., Kurbanov F., Kirpichnikov M., Protasevich I., Lobachov V., Ranjbar B., Chekhov V., Polyakov K., Engelborghs Y. and Makarov A. (1998). Comparative study of binase and barnase: experience in chimeric ribonucleases. *Protein Eng.*, **11**, 775-782.

Sheinerman F.B. and Honig B. (2002). On the role of electrostatic interactions in the design of protein-protein interfaces. *J. Mol. Biol.*, **318**, 161-177.

Shirai H., Blundell T.L. and Mizuguchi K. (2001). A novel superfamily of enzymes that catalyze the modification of guanidino groups. *Trends Biochem. Sci.*, **8**, 465-468.

Shirai H. and Mizuguchi K. (2003). Prediction of the structure and function of AstA and AstB, the first two enzymes of the arginine succinyltransferase pathway of arginine catabolism. *FEBS Lett.*, **555**, 505-510.

Shuker S.B., Hajduk P.J., Meadows R.P. and Fesik S.W. (1996). Discovering high-affinity ligands for proteins: SAR by NMR. *Science*, **274**, 1531-1534.

Sitkoff D. and Case D.A. (1998). Theories of chemical shift anisotropies in proteins and nucleic acids. *Prog. Nucl. Magn. Reson. Spec.*, **32**, 165-190.

Smith C.L., Birdsey G.M., Anthony S., Arrigoni F.I., Leiper J.M. and Vallance P. (2003). Dimethylarginine dimethylaminohydrolase activity modulates ADMA levels, VEGF expression, and cell phenotype. *Biochem. Biophys. Res. Commun.*, **308**, 984-989.

Spera S. and Bax A. (1991). Empirical correlation between protein backbone conformation and C_α and C_β ^{13}C nuclear magnetic resonance chemical shifts. *J. Am. Chem. Soc.*, **113**, 5490-5492.

Spyracopoulos L., Gagné S.M., Gronwald W., Kay L.E. and Sykes B.D. (1998). NMR studies of protein sidechain dynamics: examples from antifreeze and calcium-regulatory proteins. *Protein Dynamics, Function, and Design*, Plenum Press, NY, U.S.A., 147-162.

Stamler J.S. and Feelisch M. (1996). Preparation and detection of S-nitrosothiols. *Methods in Nitric Oxide Research,* John Wiley & Sons, Ltd, NY, U.S.A., 521-539.

Stone E.M., Person M.D., Costello N.J. and Fast W. (2005a). Characterization of a transient covalent adduct formed during dimethylarginine dimethylaminohydrolase catalysis. *Biochemistry*, **44**, 7069-7078.

Stone E.M., Schaller T.H., Bianchi H., Person M.D. and Fast W. (2005b). Inactivation of two diverse enzymes in the amidinotransferase superfamily by 2-chloroacetamidine: dimethylargininase and peptidylarginine deiminase. *Biochemistry*, **44**, 13744-13752.

Stühlinger M.C., Oka R.K., Graf E.E., Schmölzer I., Upson B.M., Kapoor O., Szuba A., Malinow M.R., Wascher T.C., Pachinger O. and Cooke J.P. (2003). Endothelial dysfunction induced by hyperhomocyst(e)inemia: role of asymmetric dimethylarginine. *Circulation*, **108**, 933-938.

Surdacki A., Nowicki M., Sandmann J., Tsikas D., Böger R.H., Bode-Böger S.M., Kruszelnicka-Kwiatkowska O., Kokot F., Dubiel J.S. and Froelich J.C. (1999). Reduced urinary excretion of nitric oxide metabolites and increased plasma levels of asymmetric dimethylarginine in men with essential hypertension. *J. Cardiovasc. Pharmacol.*, **33**, 652-658.

Tang J., Gray J.D., Clarke S. and Herschman H.R. (1998). PRMT3, a type I protein arginine N-methyltransferase that differs from PRMT1 in its oligomerization, subcellular localization, substrate specificity and regulation. *J. Biol. Chem.*, **273**, 16935-16945.

Thoma R., Hennig M., Sterner R. and Kirschner K. (2000). Structure and function of mutationally generated monomers of dimeric phosphoribosylanthranilate isomerase from *Thermotoga maritime*. *Struct. Fold. Design.*, **8**, 265-276.

Tjandra N., Feller R., Pastor R.W. and Bax A. (1995). Relaxation diffusion anisotropy of human ubiquitin from ^{15}N NMR relaxation. *J. Am. Chem. Soc.*, **117**, 12562-12566.

Tjandra N., Wingfield P., Stahl S. and Bax A. (1996). Anisotropic rotational diffusion of peudeuterated HIV protease from ^{15}N NMR relaxation measurements at two magnetic fields. *J. Biomol. NMR*, **8**, 273-284.

Tocilj A., Schrag J.D., Li Y., Schneider B.L., Reitzer L., Matte A. and Cygler M. (2005). Crystal structure of N-succinylarginine dihydrolase AstB, bound to substrate and product, an enzyme from the arginine catabolic pathway of *Escherichia coli*. *J. Biol. Chem.*, **280**, 15800-15808.

Tollinger M., Skrynnikov N.R., Mulder F.A., Forman-Kay J.D. and Kay L.E. (2001). Slow dynamics in folded and unfolded states of an SH3 domain. *J. Am. Chem. Soc.*, **123**, 11341-11352.

Tong Y., Hughes D., Placanica L. and Buck M. (2005). When monomers are preferred: a strategy for the identification and disruption of weakly oligomerized proteins. *Structure*, **13**, 7-15.

Tousignant A. and Pelletier J.N. (2004). Protein motions promote catalysis. *Chem. Biol.*, **11**, 1037-1042.

Tran C.T., Fox M.F., Vallance P. and Leiper J.M. (2000). Chromosomal localization, gene structure, and expression pattern of DDAH1: comparison with DDAH2 and implications for evolutionary origins. *Genomics*, **68**, 101-105.

Tran C.T., Leiper J. and Vallance P. (2003). The DDAH/ADMA/NOS pathway. *Atheroscler. Suppl.*, **4**, 33-40.

Traut T.W. (1994). Dissociation of enzyme oligomers: a mechanism for allosteric regulation. *Crit. Rev. Biochem. Mol. Biol.*, **29**, 125-163.

Tsai C.J., Lin S.L., Wolfson H.J. and Nussinov R. (1997). Studies of protein-protein interfaces: a statistical analysis of the hydrophobic effect. *Protein Sci.*, **6**, 53-64.

Tsikas D., Böger R.H., Sandman J., Bode-Böger S.M. and Frolich J.C. (2000a). Endogenous nitric oxide synthase inhibitors are responsible for the L-arginine paradox. *FEBS Lett.*, **478**, 1-3.

Tsikas D., Sandmann J., Savva A., Luessen P., Böger R.H., Gutzki F.M., Mayer B. and Frolich J.C. (2000b). Assessment of nitric oxide synthase activity *in vitro* and *in vivo* by gas chromatography-mass spectrometry. *J. Chromatogr. B*, **742**, 143-153.

Tugarinov V., Muhandiram R., Ayed A. and Kay, L.E. (2002). Four-dimensional NMR spectroscopy of a 723-residue protein: chemical shift assignments and secondary structure of malate synthase G. *J. Am. Chem. Soc.*, **124**, 10025-10035.

Tugarinov V., Hwang P.M. and Kay L.E. (2004). Nuclear magnetic resonance spectroscopy of high-molecular-weight proteins. *Annu. Rev. Biochem.*, **73**, 107-146.

Tugarinov V. and Kay L.E. (2005). Methyl groups as probes of structure and dynamics in NMR studies of high-molecular-weight proteins. *Chembiochem*, **6**, 1567-1577.

Tugarinov V., Choy W.Y., Orekhov V.Y. and Kay L.E. (2005). Solution NMR-derived global fold of a monomeric 82-kDa enzyme. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 622-627.

Usui M., Matsuoka H., Miyazaki H., Ueda S., Okuda S. and Imaizumi T. (1998). Increased endogenous nitric oxide synthase inhibitor in patients with congestive heart failure. *Life Sci.*, **62**, 2425-2430.

Valdar W.S. and Thornton J.M. (2001). Conservation helps to identify biologically relevant crystal contacts. *J. Mol. Biol.*, **313**, 399-416.

Valentine E.R. and Palmer A.G. III. (2004). Microsecond-to-millisecond conformational dynamics demarcate the GluR2 glutamate receptor bound to agonists glutamate, quisqualate, and AMPA. *Biochemistry*, **44**, 3410-3417.

Vallance P., Leone A., Calver A., Collier J. and Moncada S. (1992). Accumulation of an endogenous inhibitor of nitric oxide synthesis in chronic renal failure. *Lancet*, **339**, 572-575.

Vallance P. and Chan N. (2001). Endothelial function and nitric oxide: clinical relevance. *Heart*, **85**, 342-350.

Vallance P. and Leiper J. (2002). Blocking NO synthesis: how, where and why? *Nat. Rev. Drug. Discov.*, **1**, 939-950.

Vallance P. and Leiper J. (2004). Cardiovascular biology of the asymmetric dimethylarginine:dimethylarginine dimethylaminohydrolase pathway. *Arterioscler. Vasc. Biol.*, **24**, 1023-1030.

Vallance P., Bush H.D., Mok B.J., Hurtado-Guerrero R., Gill H., Rossiter S., Wilden J.D. and Caddick S. (2005). Inhibition of dimethylarginine dimethylaminohydrolase (DDAH) and arginine deiminase (ADI) by pentafluorophenyl (PFP) sulfonates. *Chem. Commun.*, **28**, 5563-5565.

Venters R.A., Farmer B.T. 2nd, Fierke C.A. and Spicer L.D. (1996). Characterizing the use of perdeuteration in NMR studies of large proteins: ^{13}C, ^{15}N and ^{1}H assignments of human carbonic anhydrase II. *J. Mol. Biol.*, **264**, 1101-16.

Vossenaar E.R., Zendman A.J., van Venrooij W.J. and Pruijn G.J. (2003). PAD, a growing family of citrullinating enzymes: genes, features and involvement in disease. *Bioessays*, **25**, 1106-1118.

Wang C., Karpowich N., Hunt J.F., Rance M. and Palmer A.G. III. (2004). Dynamics of ATP-binding cassette contribute to allosteric control, nucleotide binding and energy transduction in ABC transporters. *J. Mol. Biol.*, **342**, 525-537.

Wang J., Stieglitz K.A., Cardia J.P. and Kantrowitz E.R. (2005). Structural basis for ordered substrate binding and cooperativity in aspartate transcarbamoylase. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 8881-8886.

Wang L., Pang Y., Holder T., Brender J.R., Kurochkin A.V. and Zuiderweg E.R. (2001). Functional dynamics in the active site of the ribonuclease binase. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 7684-7689.

Wang Y., Wysocka J., Sayegh J., Lee Y.H., Perlin J.R., Leonelli L., Sonbuchner L.S., McDonald C.H., Cook R.G., Dou Y., Roeder R.G., Clarke S., Stallcup M.R., Allis C.D. and Coonrod S.A. (2004). Human PAD4 regulates histone arginine methylation levels via demethylimination. *Science*, **306**, 279-283.

Weiss V.H., McBride A.E., Soriano M.A., Filman D.J., Silver P.A. and Hogle J.M. (2000). The structure and oligomerization of the yeast arginine methyltransferase, Hmt1. *Nat. Struct. Biol.*, **7**, 1165-1171.

Wishart D.S and Sykes B.D. (1994). Chemical shifts as a tool for structure determination. *Methods Enzymol.*, **239**, 363-392.

Wittekind M. and Mueller L. (1993). HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the α-carbon and β-carbon resonances in proteins. *J. Magn. Reson. B*, **101**, 201-205.

Wong K.B., Fersht A.R. and Freund S.M. (1997). NMR ^{15}N relaxation and structural studies reveal slow conformational exchange in barstar C40/82A. *J. Mol. Biol.*, **268**, 494-511.

Xu D., Lin S.L. and Nussinov R. (1997). Protein binding versus protein folding: the role of hydrophilic bridges in protein associations. *J. Mol. Biol.*, **265**, 68-84.

Xu D., Tsai C.J. and Nussinov R. (1998). Mechanism and evolution of protein dimerization. *Protein Sci.*, **7**, 533-544.

Yamazaki T., Lee W., Arrowsmith C.H., Muhandiram D.R. and Kay L.E. (1994a). A suite of triple resonance NMR experiments for the backbone assignment of ^{15}N, ^{13}C, ^{2}H labeled proteins with high sensitivity. *J. Am. Chem. Soc.*, **116**, 11655-11666.

Yamazaki T., Lee W., Revington M., Mattiello D.L., Dahlquist F.W., Arrowsmith C.H. and Kay L.E. (1994b). An HNCA pulse scheme for the backbone assignment of ^{15}H, ^{13}C, ^{2}H-labeled proteins: Application to a 37-kDa *Trp* repressor-DNA complex. *J. Am. Chem. Soc.*, **116**, 6464-6465.

Yang D. and Kay L.E. (1999). TROSY triple-resonance four-dimensional NMR spectroscopy of a 46 ns tumbling protein. *J. Am. Chem. Soc.*, **121**, 2571-2575.

Yung A., Turnbull W.B., Kalverda A.P., Thompson G.S., Homans S.W., Kitov P. and Bundle D.R. (2003). Large-scale millisecond intersubunit dynamics in the B subunit homopentamer of the toxin derived from *Escherichia coli* O157. *J. Am. Chem. Soc.*, **125**, 13058-13062.

Zhang O., Kay L.E., Olivier J.P. and Forman-Kay J.D. (1994). Backbone ^{1}H and ^{15}N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. *J. Biomol. NMR*, **4**, 845-858.

Zhang X., Zhou L. and Cheng X. (2000). Crystal structure of the conserved core of protein arginine methyltransferase PRMT3. *EMBO J.*, **19**, 3509-3519.

Zhang X. and Cheng X. (2003). Structure of the predominant protein arginine methyltransferase PRMT1 and analysis of its binding to substrate peptides. *Structure*, **11**, 509-520.

Zoccali C., Bode-Böger S., Mallamaci F., Benedetto F., Tripepi G., Malatino L.S., Cataliotti A., Bellanuova I., Fermo I., Frolich J.C. and Böger R.H. (2001). Plasma concentration of asymmetrical dimethylarginine and mortality in patients with end-stage renal disease: a prospective study. *Lancet*, **358**, 2113-2117.

Zuiderweg E.R. (2002). Mapping protein-protein interactions in solution by NMR spectroscopy. *Biochemistry*, **41**, 1-7.